M100 9

THE OPEN UNIVERSITY

Mathematics Foundation Course Unit 9

Integration

The Open University

Mathematics Foundation Course Unit 9

INTEGRATION I

Prepared by the Mathematics Foundation Course Team

Correspondence Text 9

The Open University Press

Open University courses provide a method of study for independent learners through an integrated teaching system including textual material, radio and television programmes and short residential courses. This text is one of a series that make up the correspondence element of the Mathematics Foundation Course.

The Open University's courses represent a new system of university level education. Much of the teaching material is still in a developmental stage. Courses and course materials are, therefore, kept continually under revision. It is intended to issue regular up-dating notes as and when the need arises, and new editions will be brought out when necessary.

Further information on Open University courses may be obtained from The Admissions Office, The Open University, P.O. Box 48, Bletchley, Bucks.

The Open University Press
Walton Hall, Bletchley, Bucks

First Published 1971

Printed in Great Britain by
J W Arrowsmith Ltd., Bristol 3.

SBN 335 01008 3

Contents

		Page
	Objectives	iv
	Structural Diagram	v
	Glossary	vi
	Notation	viii
	Bibliography	viii
	Introduction	1
9.1	**The Use of Bounded Estimates in Finding Area**	3
9.1.0	What is Area?	3
9.1.1	Region Bounded by a Curve Specified by a Tabulated Function	5
9.1.2	Region Bounded by a Curve Specified by a Continuous Function	12
9.1.3	Summary	19
9.2	**The Definite Integral**	19
9.2.0	Introduction	19
9.2.1	Area and the Definite Integral	21
9.2.2	The Definite Integral for Combinations of Functions	31
9.3	**Some Applications of the Definite Integral**	39
9.3.0	Introduction	39
9.3.1	Volume of a Solid of Revolution	40
9.3.2	Averages	43
9.3.3	Velocity and Distance	45
9.3.4	Some Other Applications	46
9.4	**More Refined Approximation Methods for Definite Integrals**	47
9.4.0	Introduction	47
9.4.1	The Trapezoidal Rule	49
9.4.2	Simpson's Rule	52
9.5	**Conclusion**	55

Objectives

The primary aim of this unit is to introduce and to define precisely the concept of a definite integral.

After working through this unit you should be able to:

(i) explain the meaning of the term *definite integral*;

(ii) evaluate the definite integral of functions of the form

$$x \longmapsto ax^p + bx^q + cx^r \qquad (x \in R)$$

where a, b and c are real numbers, and p, q and r are real numbers none of which is equal to -1;

(iii) estimate the value of the definite integral for a tabulated function;

(iv) determine the upper and lower error bounds of this estimate, including the effects of errors in the tabulated values;

(v) estimate the value of the definite integral of a function using the trapezoidal rule;

(vi) estimate the number of steps (and the width of interval) required to obtain a given accuracy using this rule;

(vii) estimate the number of decimal places required in the tabulated values to achieve this accuracy;

(viii) calculate the average of a function over a given interval;

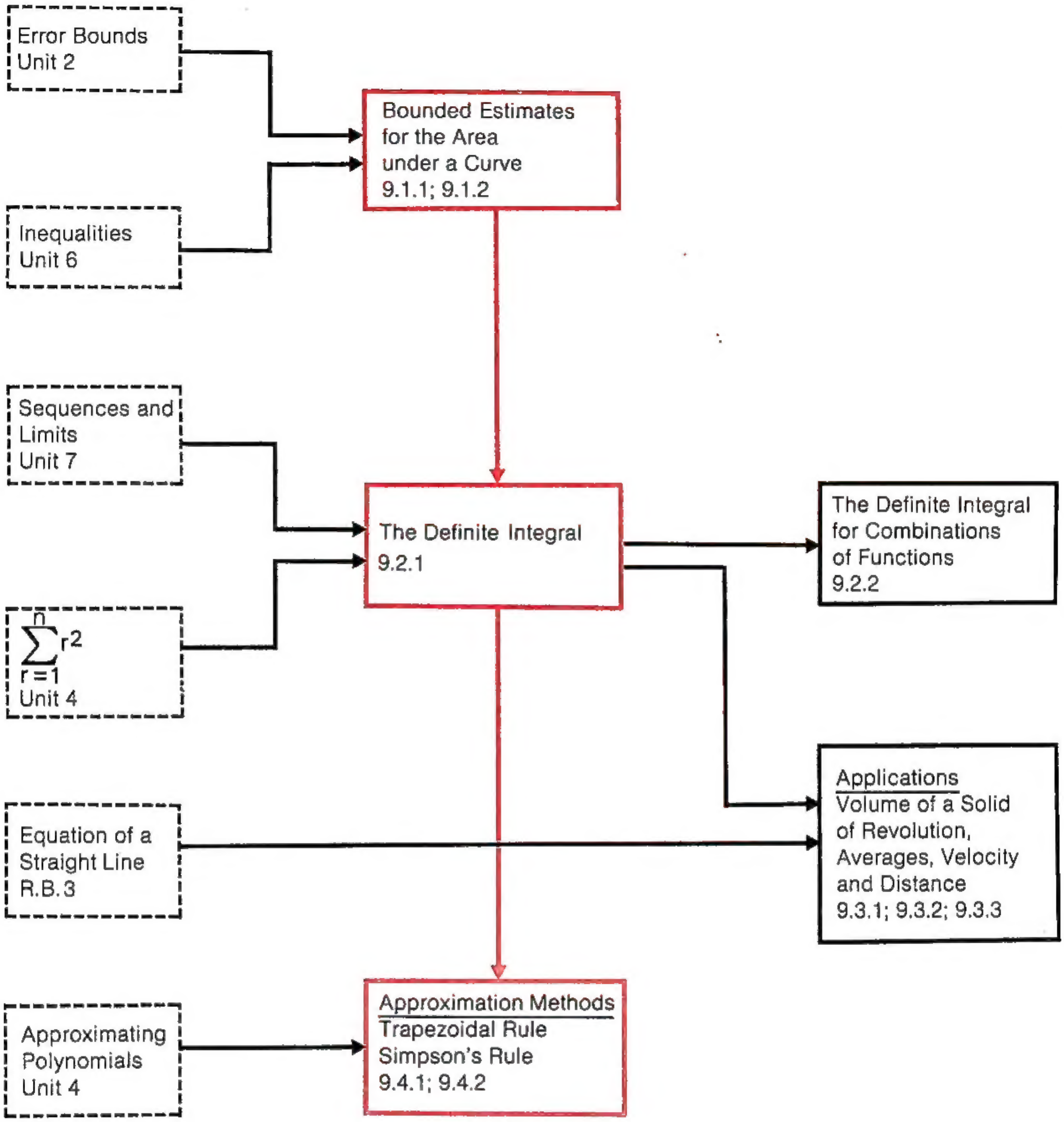
(ix) calculate the volume of revolution of a given curve;

(x) calculate the distance travelled by a body in a given time, given the velocity as a function of time;

(xi) translate the traditional Leibniz notation for a definite integral into the functional notation used in this course, and vice versa.

N.B.
Before working through this correspondence text, make sure you have read the general introduction to the mathematics course in the Study Guide, as this explains the philosophy underlying the whole course. You should also be familiar with the section which explains how a text is constructed and the meanings attached to the stars and other symbols in the margin, as this will help you to find your way through the text.

Structural Diagram



Glossary

Terms which are defined in this glossary are printed in CAPITALS.

		Page
ABSOLUTE ERROR	The ABSOLUTE ERROR in a measurement x is the difference $x - X$ between the measured number x and the exact number X.	7
ABSOLUTE ERROR BOUND	The ABSOLUTE ERROR BOUND in a measurement is the maximum possible value of the magnitude of the ABSOLUTE ERROR.	7
AREA OF A RECTILINEAR FIGURE	The area of a rectangle is base × height. The area of a right-angled triangle is $\frac{1}{2}$ × base × height. The area of a RECTILINEAR FIGURE can be deduced from these definitions.	3
AREA BENEATH THE GRAPH OF f BETWEEN a AND b	The AREA BENEATH THE GRAPH OF f BETWEEN a AND b is the limit of the sequence defined by $S_n = h[\lvert f(a)\rvert + \lvert f(a+h)\rvert + \cdots + \lvert f(a + \{n-1\}h)\rvert]$, where $h = \dfrac{b-a}{n}$, if this limit exists.	21
AVERAGE	The AVERAGE of $f(x)$ over the interval $[a, b]$ is H, where $$H = \frac{1}{b-a}\int_a^b f$$	45
BOUND	The number B is an UPPER BOUND of a set of numbers S if $s \leqslant B$ for each element $s \in S$. The number b is a LOWER BOUND of a set of numbers S if $s \geqslant b$ for each element $s \in S$. A set is BOUNDED if it has upper and lower bounds.	4
DEFINITE INTEGRAL	The DEFINITE INTEGRAL of the function f between a and b is the limit of the sequence defined by $S_n = h[f(a) + f(a+h) + \cdots + f(a + \{n-1\}h)]$ where $h = \dfrac{b-a}{n}$, if this limit exists.	24
END-POINTS OF A DEFINITE INTEGRAL	The END-POINTS OF THE DEFINITE INTEGRAL of the function f over the interval $[a, b]$ are a and b. The UPPER END-POINT is b, and the LOWER END-POINT is a.	25
INTEGRATE	To INTEGRATE means to evaluate a DEFINITE INTEGRAL.	47
ORDINATE	The ORDINATE of a point in the Cartesian plane is the y-co-ordinate of the point.	17
RECTILINEAR FIGURE	A RECTILINEAR FIGURE is a figure bounded by straight lines.	3
SIMPSON'S RULE	SIMPSON'S RULE is an approximation to the DEFINITE INTEGRAL OF f BETWEEN a AND b given by the formula $$\int_a^b f \simeq \frac{h}{3}(y_0 + 4y_1 + 2y_2 + \cdots + 4y_{n-1} + y_n)$$ where the interval $[a, b]$ is divided into n (n even) SUB-INTERVALS and $y_0, y_1, \ldots, y_n$ are the ORDINATES at the ends of these sub-intervals.	54
SUB-INTERVAL	A SUB-INTERVAL of $[a, b]$ is an interval $[c, d]$ where $a \leqslant c < d \leqslant b$.	23
TRAPEZIUM	A TRAPEZIUM is a figure with four straight sides, of which two are parallel.	4

		Page
TRAPEZOIDAL RULE	The TRAPEZOIDAL RULE is an approximation to the DEFINITE INTEGRAL OF f BETWEEN a AND b given by the formula $$\int_a^b f \simeq \frac{h}{2}(y_0 + 2y_1 + 2y_2 + \cdots + 2y_{n-1} + y_n)$$ where the interval $[a, b]$ is divided into n SUB-INTERVALS and $y_0, y_1, \ldots, y_n$ are the ORDINATES at the ends of these sub-intervals.	50
VOLUME OF REVOLUTION	A VOLUME OF REVOLUTION is the volume of a solid bounded by a surface, the surface being generated by rotating the graph of a function f in an interval $[a, b]$ about the x-axis.	40
WIDTH OF AN INTERVAL	The WIDTH OF THE INTERVAL $[a, b]$ is $b - a$.	23

Notation

The symbols are presented in the order in which they appear in the text.

		Page
e_x	The absolute error in a measurement x.	7
ε	The absolute error bound.	15
a_n	The sum of the areas of n rectangles, which gives an estimate of a required area which is less than the required area.	21
$\underset{\sim}{a}$	The sequence $a_1, a_2, a_3, \ldots$	21
A_n	The sum of the areas of n rectangles, which gives an estimate of a required area which is greater than the required area.	
$\underset{\sim}{A}$	The sequence $A_1, A_2, A_3, \ldots$	21
A	The common limit of $\underset{\sim}{a}$ and $\underset{\sim}{A}$.	23
h	The width of a sub-interval of $[a, b]$ when $[a, b]$ is divided into n equal sub-intervals. That is, $h = \dfrac{b-a}{n} \quad (n \in Z^+)$	23
S_n	The sum $h[f(a) + f(a+h) + \cdots + f(a + \{n-1\}h)]$.	23
$\int_a^b f$	The definite integral of f in $[a, b]$.	24

Bibliography

Two books which introduce integration before differentiation are mentioned below. The approach to the definite integral in both is more rigorous than the approach in our course, and would be useful reading for those who wish to extend themselves at this stage. However, do not be surprised if you find them a little difficult.

T. M. Apostol, *Calculus Vol. I* (Blaisdell 1967). See Part 1 of the Introduction, and also Section 1.

B. Hunt, *Calculus and Linear Algebra* (W. H. Freeman 1967). See Chapters 8 and 9.

9.0

Introduction

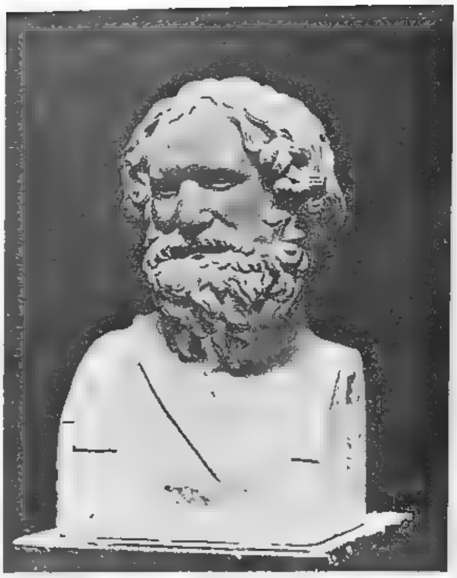
9.0 INTRODUCTION

The definite integral is a mathematical concept which arose from the problem of determining area. This problem links together two major ideas from previous units — *inequalities* and *limiting processes* — to produce a mathematical formulation of the meaning of area. From this we obtain the concept of the definite integral, which we find has many other applications.

If you have met the topic of integration before, as well as the topic of differentiation which appears later in our course, you may be wondering why we introduce the definite integral first. It is perfectly possible to introduce these topics in either order, and each approach can be made logically consistent. In fact, we shall introduce integration and differentiation each in its own right, to emphasize that each is an important mathematical concept. In a later unit, we shall establish the connection between them in the Fundamental Theorem of Calculus. We begin with the definite integral because historically the basic ideas of this were known long before those of differentiation, although the rigorous treatment of the integral came at a much later stage.

If you have some familiarity with the calculus you may also wonder why here, and in the later units on integration and differentiation, we use a notation which is different from the classical notation with which you are familiar, and which is used in most textbooks. (In fact, you will find that there are a number of recent textbooks which use the same notation as ours.) There are several reasons for this change. The notation we use is consistent with our basic approach to mathematics through the concept of a function. Also, we wish to avoid some of the conceptual difficulties experienced by many students, which may in part be due to the traditional notation. If you are new to the subject there should be no notational difficulty. If you have already studied calculus, we suggest that you make a fresh start: the new notation should enable you to concentrate on the principles (as opposed to the techniques), because of its very novelty. Once you have mastered the basic principles, there is no objection to your using the traditional notation: the conversion from one notation to the other should not prove difficult.

The history of the problem of finding area is an interesting one. In the early years of ancient Babylon it was believed that the area of a plane figure depended on its perimeter. (If you were an Estate Agent in Babylon, how would you buy and sell your land to take as much advantage of this belief as possible?)



Archimedes (Mansell)

However, the correct methods for finding areas of rectangles and triangles were known before 2200 B.C. The next step, that of finding areas of plane figures with specific curved boundaries, such as a parabolic arc, was apparently not taken until the time of Archimedes (287–212 B.C.). His mathematical method is essentially the same as the one we develop in this unit. One of his elementary checks was to cut out the appropriate shape in a material of uniform density, and compare its weight with that of a shape of the same material and of known area. Archimedes realized that the problem of the determination of volumes bounded by curved surfaces was similar to the problem of determination of area, and that both could be approached by using a process involving closer and closer approximations. Much later, in the seventeenth century, Newton, amongst others, formalized integration and established its link with differentiation.

Newton's work aroused enormous interest, and the names of many mathematicians of renown appear in the history of the calculus. By the time Riemann published his definition of the definite integral in 1854, mathematicians had long realized that they had in their hands an extremely powerful tool.

In this unit we develop the theory of integration along the following lines. We first consider ways of estimating area, and we note that elementary methods of calculating areas precisely can only cope with regions bounded by straight lines. Other regions require a different approach: we approximate to the region by successive sets of rectangles (or other suitable shapes of known area), and thus obtain a sequence of what we intuitively consider to be better and better approximations to the area. The limit of this sequence of approximations is called a *definite integral.* This concept can be applied to a far wider range of problems than the calculation of areas. For example, integration is used in determining volumes, and rate of flow of a liquid through a pipe, as you will see in the television programme associated with this unit.

We shall see later that, even though we can represent an area by a definite integral, there are cases when we cannot evaluate it in practice. In these circumstances we are forced to return to *estimating* the area; we conclude the unit with a short discussion of two methods of doing this.

You may have gathered from what we have been saying that we are now going to tackle this problem with the guns we have been priming in earlier units. We shall, in particular, be using the ideas of *function, accuracy and errors*, and *limit of an infinite sequence.* The next few units should help to show the importance of some of the ideas in the earlier units.

9.1 THE USE OF BOUNDED ESTIMATES IN FINDING AREA

9.1

9.1.0 What is Area?

9.1.0

Discussion

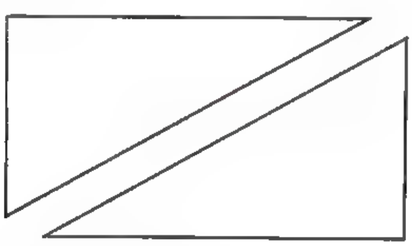
Intuitively we all know what we mean by area, just as we feel we know the meanings of length, time, speed and volume. We are also aware that we use each of these words in two senses: sometimes to mean a physical quantity, and sometimes to mean a measurement of that quantity. For example, the word "area" means a vacant piece of level ground; but to avoid using a clumsy phrase, we say "the area is five acres", where the word "area" stands for "the measurement of the area". The sense intended is usually obvious from the context. Here, we are interested in the measurement of area. We can leave the philosophers to worry about the fundamental concept. As mathematicians, we use our intuition to guide us to a precise mathematical formulation of area; we then develop and generalize the mathematical concept we have defined. You have already experienced such a development in *Unit 7, Sequences and Limits I*, where we discussed a possible mathematical formulation of the intuitive concept of velocity at an instant.

Definition 1

Using our intuition, let us *define** the area of a rectangle as the product of the lengths of two adjacent sides. With this as our starting point, we shall proceed to define areas of other figures of increasing complexity, always assuring ourselves that our definitions accord with our intuitive expectations.

Definition 2

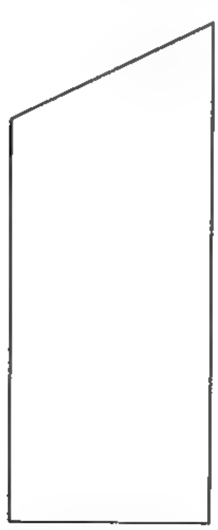
Let us start with rectilinear figures, which are figures bounded by straight lines. By fitting together two identical right-angled triangles to form a rectangle, we find that the area of such a triangle is half the product of the length of a base and the corresponding height, which we shall write as $\frac{1}{2} \times$ base $\times$ height.



Knowing this, we can find the area of any rectilinear figure by regarding the figure as being built up of rectangles and right-angled triangles.

Exercise 1

Exercise 1
(2 minutes)



* We define area such that it has dimensions L^2 (see *Unit 3, Operations and Morphisms*, page 36). We shall state the units of area when they are known: otherwise we shall only give the magnitude of the area.

A trapezium is a quadrilateral which has two parallel sides. Definition 3

Show that the area of the particular trapezium shown is $\frac{1}{2}ad$, where

a = the sum of the lengths of the parallel sides, and

d = the perpendicular distance between them. ■

Discussion

Difficulties arise however, when we consider the area of a region bounded by a curve such as a circle or an arch. We can, like Archimedes, cut out the appropriate shape and weigh it, and then say that the area is the area of the rectangle of the same material and weight. A mathematical definition is more complicated and is phrased in terms of limiting processes, but it is much more satisfactory.

In the next section we shall use numerical data to illustrate how we can *estimate* an area (and a volume) in a practical case. The purpose of this is twofold. Firstly, we shall see how we can use the areas of rectangles to determine upper and lower values (called bounds) between which we intuitively know the area to lie. Secondly we shall show that, when *actual* measurements are used in an estimation, there is no need to involve any limiting processes, since no increase in accuracy is thus obtained. Therefore, unless you are particularly interested in this part of the text, do not get too involved with the arithmetic; read through it quite quickly and grasp the two ideas mentioned above.

We shall begin by considering the practical case when part of the boundary of the shape whose area we wish to find is specified by a tabulated function, where the tabular points are equally spaced.

We consider a vertical cross-section of a construction site; that is, the area we consider represents an imaginary straight slice down through the ground. A part of the boundary of such a cross-section may be specified approximately by tabulated values of the heights above a reference level of equally spaced points on the boundary.

We shall see that the accuracy of the estimate of an area which we find by taking observations, is usually restricted by the accuracy of the observations rather than by the estimation method used.

We encounter no similar hindrance to accuracy in the calculation of area when the bounding curve is the graph of a function such as

$$f: x \longmapsto 10 - \frac{x^2}{10} \qquad (x \in R)$$

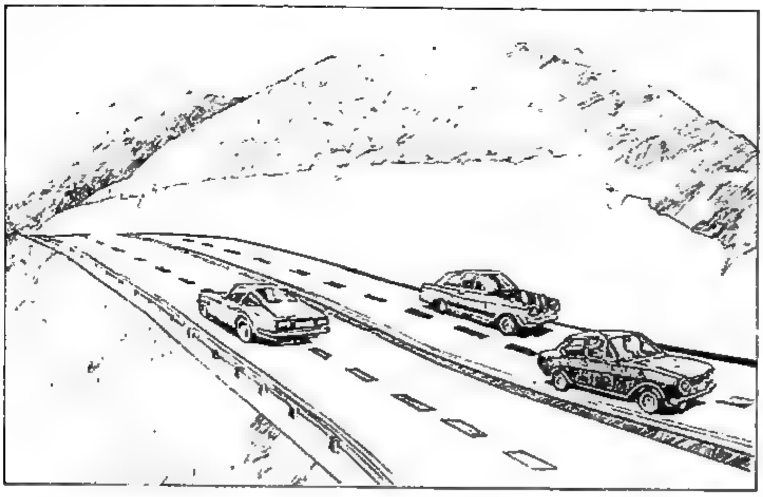
We shall see that we can then estimate the area as accurately as we wish. This will enable us to define the area in a mathematical sense in section 9.2.

To sum up: in section 9.1 we shall use our knowledge of the areas of rectangles to determine other areas as accurately as we can.

9.1.1 Region Bounded by a Curve Specified by a Tabulated Function

9.1.1

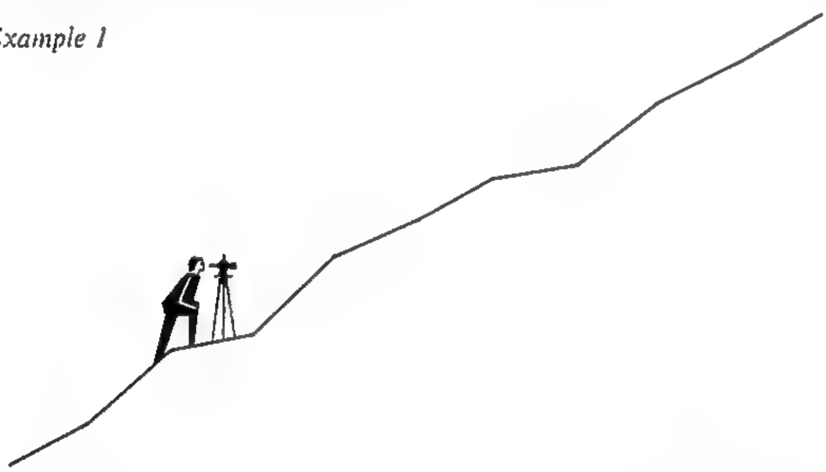
Discussion



The sketch above shows a motorway skirting the side of a hill. To construct this section of the motorway, a large amount of soil and rock would have to be removed. To determine the volume of material to be transported, any contractor would in practice have to take into account the fact that the cross-section of the volume to be removed varies along the direction of the road itself, but for illustrative purposes we shall consider the problem of determining the volume when the area of the cross-section is assumed constant.

Example 1

Example 1

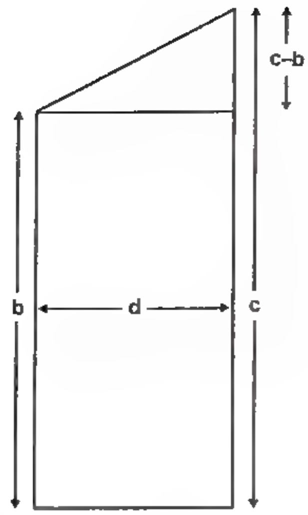


As a result of carrying out a survey of the site, the following measurements were found for the cross-sectional area (of width 20 m) used to determine the volume of material to be removed:

Horizontal distance in metres	Height above reference level in metres
0	0
2	1.3
4	3.4
6	3.9
8	6.1
10	7.2
12	8.5
14	8.9
16	10.7
18	11.8
20	13.3

(*continued on page 6*)

Solution 1



In the above diagram,

$$\text{the area of the triangle} = \tfrac{1}{2}d(c - b)$$

$$\text{the area of the rectangle} = db$$

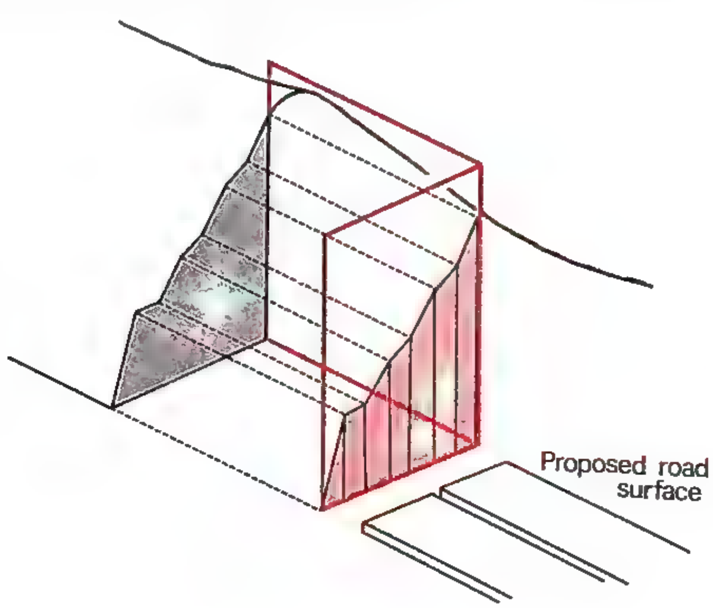
So

$$\begin{aligned}\text{the area of the trapezium} &= db + \tfrac{1}{2}dc - \tfrac{1}{2}db \\ &= \tfrac{1}{2}(c + b)d \\ &= \tfrac{1}{2}ad\end{aligned}$$

■

(*continued from page 5*)

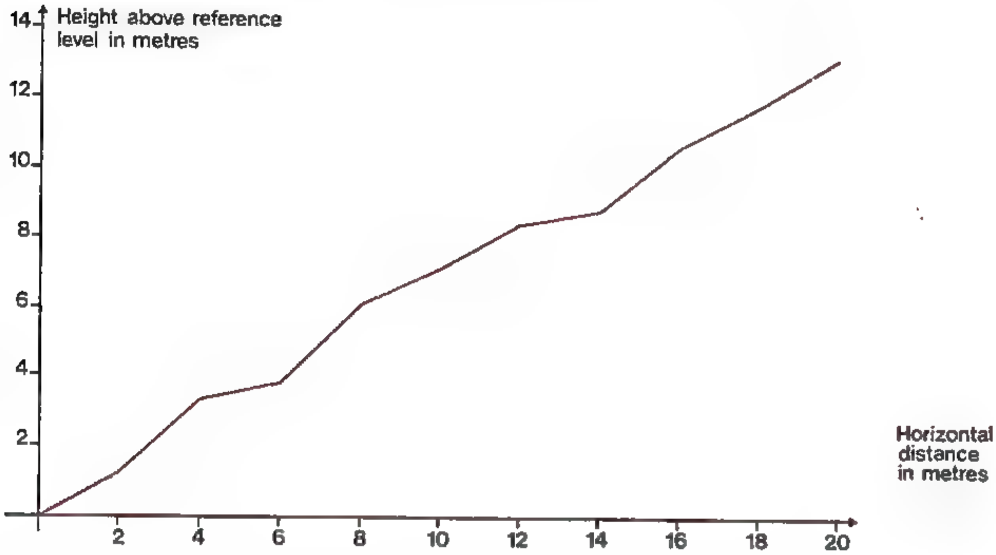
There is a limit to the accuracy with which we can take the measurements. Let us assume that there is an error of ± 0.05 m in all the horizontal steps and in all the heights. (Cumulative effects of such errors could make things worse — this is the *best* we can do.) The problem is to determine, as accurately as we can, the cross-sectional area of the material to be removed in this simplified case; the volume can then be determined by multiplying by the length of the stretch of road.



■

Solution of Example 1

Remember the elementary way we are going to approach this problem. We are going to *approximate* the area by a sum of the areas of rectangles. But it would also be desirable to obtain an estimate of the magnitude of the error we may be making. We can do this fairly simply by determining two areas made up of rectangles: one which will be larger than the area we require and one which will be smaller. We then know upper and lower bounds of the area we are interested in. We can plot the data on a graph like the one given below.



Put Overlay 1* in position. An extremely crude guess at the answer is to say that the area lies somewhere between the area of the rectangle *ABCD* and zero. But we can do better than this. Put Overlay 2 in position. You then see an area, made up of rectangles, which is larger than the area we require. We have used the heights at 4 m (instead of 2 m) steps in our second attempt to estimate the area, in order to demonstrate later the improvement we get in accuracy as we increase the number of rectangles appropriately. You may well ask at this stage how we know that the boundary of the area obtained using rectangles encloses the actual surface, since we only have a finite number of observations and there might be a sharp peak in the ground in between two of these observations. We do *not* know this, of course, but, as in finite differences, when we have the tabulated values of a function we assume that the graph of the function in between the tabulated points is reasonably smooth. (Anyhow, the surveyor should have spotted any irregularities.)

The area of the sum of the rectangles on Overlay 2 is 168.0 m². We now estimate the absolute error bound in this area, due to the absolute error bound of 0.05 m in the measurements of height and of the horizontal step length.

We use e_x to denote the absolute error in the measurement x. In *Unit 2, Errors and Accuracy*, section 2.3, we deduced that

$$e_{x+y} = e_x + e_y$$

and

$$e_{xy} = ye_x + xe_y - e_xe_y$$

that is,

$$e_{xy} \simeq ye_x + xe_y$$

since the term e_xe_y is small.

* The overlays are in the wallet on the inside of the back cover of this text.

Applying the last result to find the error in the area of each of our rectangles, and then applying the first result to the sum of these areas, we get an estimated absolute error in the total area of all the rectangles:

$$(0.05(3.4 + 6.1 + 8.5 + 10.7 + 13.3) + 0.05(4 + 4 + 4 + 4 + 4))\,\mathrm{m}^2 = 3.1\,\mathrm{m}^2$$

since the absolute error bound in each of our measurements is 0.05 m.

Thus the estimated absolute error bound in the area is $3.1\,\mathrm{m}^2$.

Now put Overlay 3 on the basic graph. The area shown here is clearly less than the area we need (we assume that there are no deep, narrow ditches in the ground).

The sum of the areas of the rectangles is now $114.8\,\mathrm{m}^2$, with estimated absolute error bound:

$$(0.05(3.4 + 6.1 + 8.5 + 10.7) + 0.05(4 + 4 + 4 + 4))\,\mathrm{m}^2 = 2.2\,\mathrm{m}^2$$

Can we improve this accuracy by using smaller rectangles, and in particular, by using rectangles with half the width of the previous ones?

Put Overlay 4 on the basic graph. You will now see two sets of rectangles on a 2 m wide base. If you use Overlays 2 and 3 at the same time as Overlay 4, you will see that the sum of the areas of the larger rectangles with bases 2 m wide is less than the sum of the areas of the larger rectangles with bases 4 m wide, but it is still greater than the required area. Similarly the sum of the areas of the smaller rectangles with bases 2 m wide is greater than the corresponding sum for intervals of 4 m, but it is still less than the required area.

Carrying out the calculations for this last case we get for the larger area:

$$2(1.3 + 3.4 + 3.9 + 6.1 + 7.2 + 8.5 + 8.9 + 10.7 + 11.8 + 13.3)\,\mathrm{m}^2 = 150.2\,\mathrm{m}^2$$

with an estimated absolute error bound of $4.8\,\mathrm{m}^2$*

Similarly, for the smaller area:

$$2(1.3 + 3.4 + 3.9 + 6.1 + 7.2 + 8.5 + 8.9 + 10.7 + 11.8)\,\mathrm{m}^2 = 123.6\,\mathrm{m}^2$$

with an estimated absolute error bound of $4.0\,\mathrm{m}^2$.*

Note that, in this example, the difference between the larger and smaller areas is simply the area of the largest rectangle. Can you see why this is so? This simplifies the calculation, since we can calculate the larger area from the smaller area by adding to it the area of the largest rectangle. See T.V.

Thus the area has a value between

$$(123.6 - 4.0)\,\mathrm{m}^2 = 119.6\,\mathrm{m}^2$$

and

$$(150.2 + 4.8)\,\mathrm{m}^2 = 155.0\,\mathrm{m}^2$$

and we have apparently narrowed the gap between our upper and lower estimates by increasing the number of intervals we use on the distance-axis from five to ten. The following exercises develop this idea a little further. The first of these exercises indicates a method of obtaining the "best" estimate of a number when we know the error interval in which it is contained.

■

* If you check these results, remember that we are assuming that each height and each step length has an absolute error bound of 0.05 m.

Exercise 1

Exercise 1
(2 minutes)

If you were told that the number of matches in a box was between 44 and 50 inclusive, how many matches would you say were in a specific box if your objective was to make the possible error a minimum?

What would be the absolute error bound in this case? ■

Exercise 2

Exercise 2
(3 minutes)

(i) Using the results (obtained on pages 7 and 8) for five intervals in our example, and the idea introduced in the last exercise, give an estimate of the cross-sectional area in Example 1. State the estimated absolute error bound due to the propagation of errors from the inexact data, and the absolute error bound arising from the use of rectangular approximations.

(ii) Repeat (i) using the results on page 8 for ten intervals. ■

Discussion

As we increase the number of rectangles, the estimated absolute error bound arising from the inexact data will continue to increase, while that arising from the rectangle approximation will decrease (as we saw in the two parts of the solution to the last exercise). After a certain stage, this makes the attempt to further improve accuracy by using more and more intervals a worthless pursuit.

Exercise 3

Exercise 3
(3 minutes)

Describe in words, with the aid of sketches, how you would determine an estimate with upper and lower bounds of the *volume* of material removed if the cross-sectional area changed significantly along the line of the road.

(HINT: Think what to use for your basic element of volume.) ■

Solution 1

Solution 1

Forty-seven matches with a possible error of 3. If you choose a value other than 47, say 46, your error could be greater, for example, it would be 4 if there were 50 matches in the box. So the answer is $\frac{1}{2}(50 + 44)$ matches, and the absolute error bound is $\frac{1}{2}(50 - 44)$. ■

Solution 2

Solution 2

(i) Largest possible area $= 168.0 + 3.1 = 171.1\ \text{m}^2$

Smallest possible area $= 114.8 - 2.2 = 112.6\ \text{m}^2$

Estimated area $= \frac{1}{2}(171.1 + 112.6) = 141.8\ \text{m}^2$

Estimated absolute error bound $= \frac{1}{2}(171.1 - 112.6) = 29.2\ \text{m}^2$

of which:

estimated absolute error bound arising from inexact data $= \frac{1}{2}(3.1 + 2.2) = 2.6\ \text{m}^2$

and

absolute error bound arising from the rectangle approximation $= \frac{1}{2}(168.0 - 114.8) = 26.6\ \text{m}^2$

(ii) Largest possible area $= 150.2 + 4.8 = 155.0\ \text{m}^2$

Smallest possible area $= 123.6 - 4.0 = 119.6\ \text{m}^2$

Estimated area $= \frac{1}{2}(155.0 + 119.6) = 137.3\ \text{m}^2$

Estimated absolute error bound $= \frac{1}{2}(155.0 - 119.6) = 17.7\ \text{m}^2$

of which:

estimated absolute error bound arising from inexact data $= \frac{1}{2}(4.8 + 4.0) = 4.4\ \text{m}^2$

and

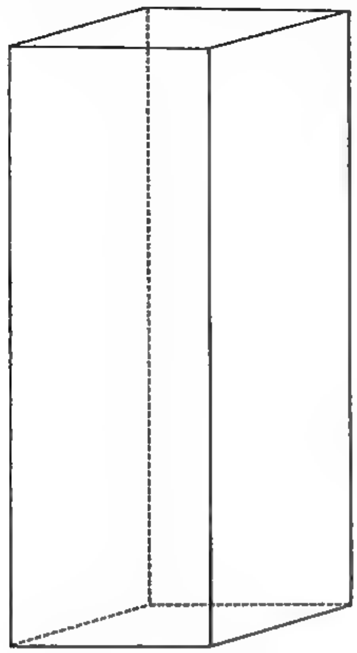
absolute error bound arising from the rectangle approximation $= \frac{1}{2}(150.2 - 123.6) = 13.3\ \text{m}^2$

which is exactly half that in (i), as expected. ■

Solution 3

Solution 3

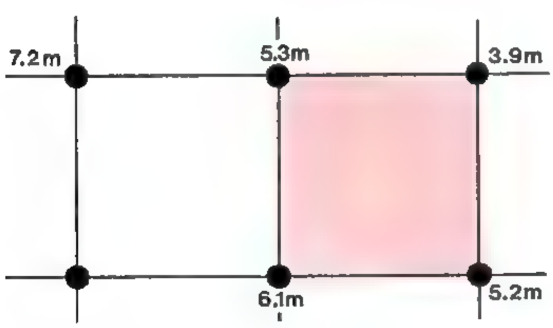
A possible method is outlined below, using a basic element of volume:



(i) The element of volume is a rectangular prism on a square base. We define the volume of such a prism to be the product of the area of the base and the height.

Definition 1

(ii) A grid of heights is plotted on the surface with equal distances in both horizontal directions, for example:



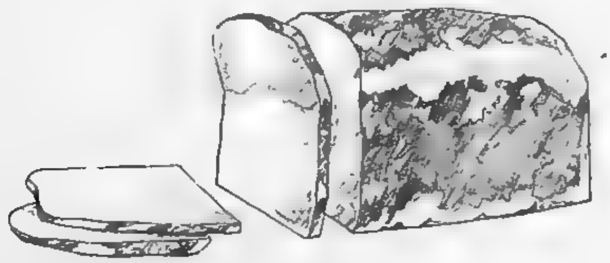
(iii) Take the maximum height in each grid square, for example, 6.1 m in the one shaded. This height will lead to a rectangular prism based on this square with volume *greater* than the volume required.

(iv) Take the minimum height in the grid, for example, 3.9 m. This gives the smallest volume of a corresponding prism.

(v) The volume lies between the sums of the minimum and maximum volumes of the basic rectangular prisms. The "best" estimate is half the sum of the maximum and minimum volumes.

Alternatively, you could have estimated the cross-sectional areas of slices of the site at equal intervals, like slices of a loaf of bread, and then found the maximum and minimum values of the volume of each slice.

This method has the disadvantage that you need to estimate the magnitudes of all the cross-sectional areas first.



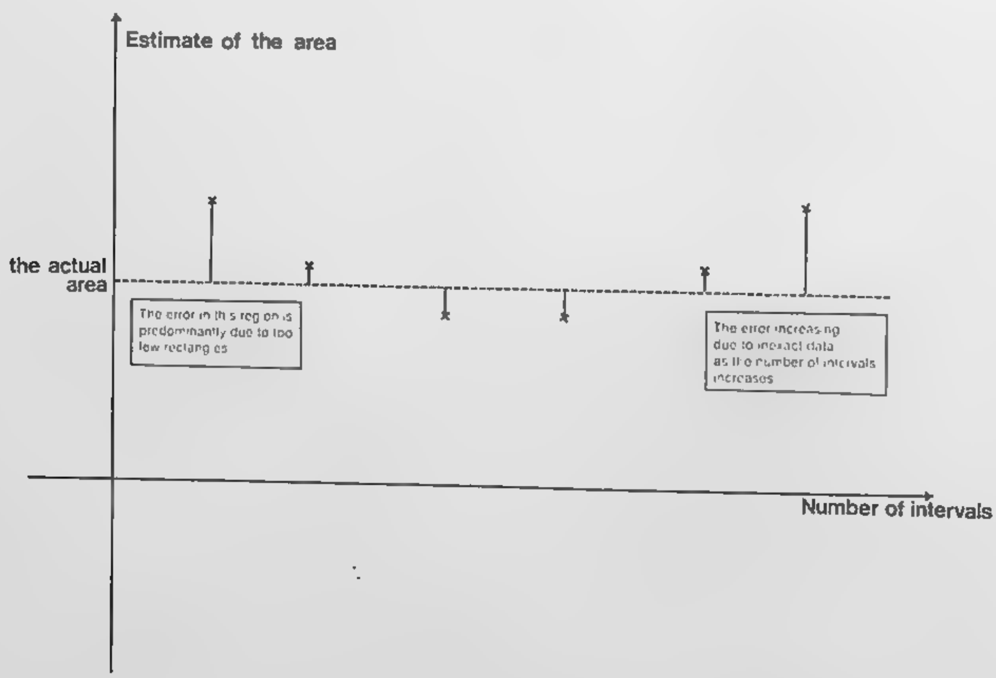
■

9.1.2 Region Bounded by a Curve Specified by a Continuous Function

9.1.2

Discussion
* *

In the previous section we saw that, where there were possible errors in the measurement of a profile, our attempt to continue to improve the accuracy of our estimated area by increasing the number of rectangles used in the calculation was thwarted. It seemed that, if we reduced the width of the intervals, we got a better approximation to the area *until* the inherent errors began to dominate.



The crosses on the diagram represent a plausible set of calculated values for the area (not the actual values) as the number of intervals is increased. For a small number of intervals the error is dominated by the error in the rectangle approximation, for a large number of intervals by the inherent errors in the data. In the example in the last section, the figures on page 10 suggest that there would be little point in reducing the horizontal steps to much less than 1 m. In the extreme case, if we can only measure the width of the intervals to within 0.05 m, there is obviously

not much point in reducing the interval width to less than 0.05 m. (In fact, as we have seen, the reduction of the interval width would lose its point at a much larger value.)

If the profile which determines the boundary of the area is given exactly at *each* point, then the last restriction does not apply. This would be the case if the profile were the graph of the function

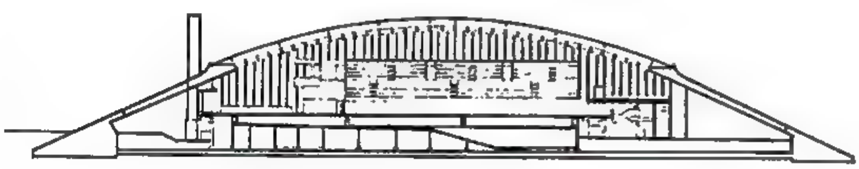
$$f : x \longmapsto 10 - \frac{x^2}{10} \qquad (x \in [-10, 10])$$

as in the next example. This is the sort of situation which would arise if we were designing a piece of equipment or a building, where *we* decide what shapes we want, as opposed to the previous example, where the shape was provided for us and we had to do the best we could.

For a region bounded by a curve specified by a continuous function, we can, by dividing the domain of the function into as many intervals as we please, approximate the area to whatever accuracy we desire. Does this stir memories of the limiting procedures and the concept of continuity of a function introduced in *Unit 7, Sequences and Limits I*? We hope so, because this is the way we shall define the area as a limit in section 9.2.1. Let us first look at a specific example:

Example 1 **Example 1**

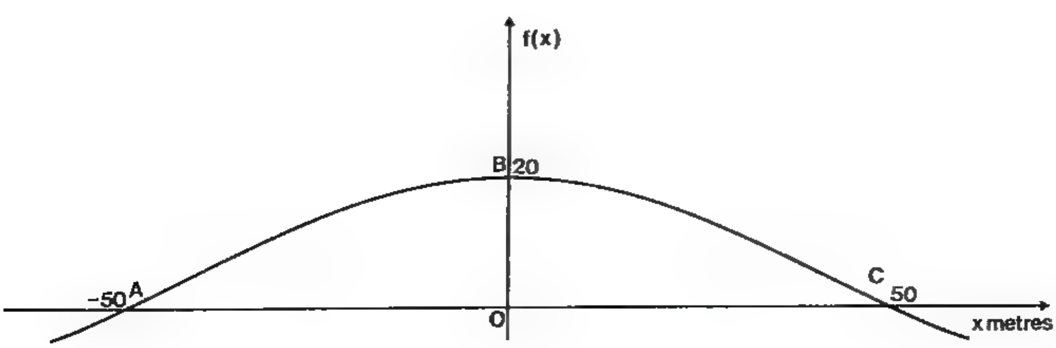
The Dollan Baths, Scotland's Olympic length swimming baths, were officially opened on May 27, 1968, by R. B. McGregor, the Scottish International swimmer.



Designed by architect A. Buchanan Campbell, the building rises in one immense parabolic arch, which has a span of 100 metres and which reaches a maximum height of 20 metres above ground level.

When he had decided on the shape, the architect had to calculate the cross-sectional area of the building to enable him to find the pressure exerted on the structure.

In order to emulate his calculation, we first need to find a function f whose graph is a curve which represents the outline of the building on a suitable scale.



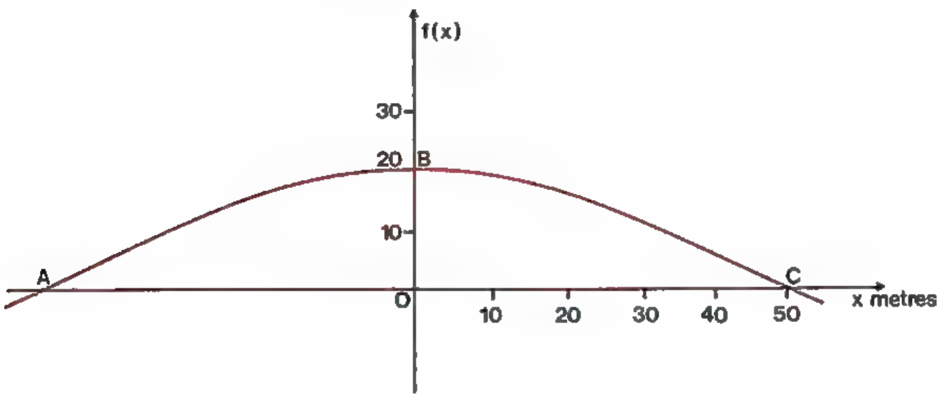
You may like to try to find this function, knowing that its graph is part of a parabola, which means that it is a function of the form

$$x \longmapsto a + bx + cx^2 \qquad (a, b, c \in R)$$

You should get the answer $x \longmapsto 20 - \dfrac{x^2}{125}$. However, our aim is to calculate the area underneath this graph. There are two parts to this:

(i) we approximate to the area with rectangles, find upper and lower estimates of the cross-sectional area of the parabola and hence find the best estimate of the area in the two cases:
   (a) when OC is divided into five equal intervals;
   (b) when OC is divided into ten equal intervals;

(ii) we determine how many intervals of OC would be required to ensure that the estimated absolute error bound is less than 1 m^2, assuming that exact arithmetic is used.



■

Solution of Example 1

(i) Have a look at the basic graph without any of the overlays* in place. The parabolic arch is symmetric; therefore, the area represented by $AOCB$ is twice the area represented by OBC. The images for the given values of the variable x are calculated exactly:

x	0	5	10	15	20	25	30	35	40	45	50
$f(x)$	20	19.8	19.2	18.2	16.8	15	12.8	10.2	7.2	3.8	0

(a) Now put on Overlay 5.

Using five intervals for the half-area, form the area consisting of rectangles which give an area larger than the area required. These rectangles are outlined by solid lines in the figure. This gives us a first upper estimate for the *total* area:

$$2 \times 10\{20 + 19.2 + 16.8 + 12.8 + 7.2\}\ \text{m}^2 = 1520\ \text{m}^2$$

The dashed lines in the figure indicate an area made up of rectangles which give an area smaller than the area we require.

This gives a first lower estimate for the total area:

$$2 \times 10\{19.2 + 16.8 + 12.8 + 7.2 + 0\}\ \text{m}^2 = 1120\ \text{m}^2$$

We notice that the difference between the two estimates is twice the area of the largest rectangle.

* The overlays are in the wallet on the inside of the back cover of this text.

Thus the best estimate for the total area using five intervals is

$$\frac{1520 + 1120}{2}\,\mathrm{m}^2 = 1320\,\mathrm{m}^2$$

with estimated absolute error bound

$$\frac{1520 - 1120}{2}\,\mathrm{m}^2 = 200\,\mathrm{m}^2$$

(b) Remove Overlay 5 and put on Overlay 6.

When ten intervals are used for the half area, we find that the upper bound on the area using ten rectangles is now

$$2 \times 5\{20 + 19.8 + \cdots + 7.2 + 3.8\}\,\mathrm{m}^2 \qquad = 1430\,\mathrm{m}^2$$

Calculation of the lower bound gives an area $= 1230\,\mathrm{m}^2$

Our estimate would be $\left(\dfrac{1430 + 1230}{2}\right)\mathrm{m}^2 \quad = 1330\,\mathrm{m}^2$

with estimated absolute error bound now $= 100\,\mathrm{m}^2$

(ii) We are now looking at a different aspect of the problem.

We state at the beginning how accurate we want our estimate of the total area to be. We want it to be accurate to within $\pm 1\,\mathrm{m}^2$, that is, we require an absolute error bound of $1\,\mathrm{m}^2$.

For what width of interval can we achieve this? As we have already noticed, the difference between the upper and lower bounds of the estimate of the area is twice the area of the largest rectangle, and it is also twice the estimated absolute error bound. Therefore the area of this rectangle must be less than or equal to $1\,\mathrm{m}^2$. But in this example we know that the height of the largest rectangle is equal to 20 m regardless of the number of intervals. Therefore, to achieve the desired accuracy, its width must be at most 0.05 m, since

$$20 \times 0.05\,\mathrm{m}^2 = 1\,\mathrm{m}^2$$

The total number of intervals along $OC = \dfrac{50\,\mathrm{m}}{\text{interval width}} = 1000$. So to be sure of achieving the desired accuracy by this method, we must divide OC into at least 1000 intervals.

Similarly, we can show that, to achieve an estimate of the total area with absolute error bound $0.1\,\mathrm{m}^2$, we would require at least 10 000 intervals; and that in general to achieve an estimate with absolute error bound $\varepsilon\,\mathrm{m}^2$ we would require at least n sub-intervals where n is an integer such that

$$n \geqslant \frac{1000}{\varepsilon}$$

N.B. Later in this text we shall see that we can find the above parabolic area exactly. Its value is $1333\frac{1}{3}\,\mathrm{m}^2$. So, in fact, the above estimates are very good ones with an error which is already less than $15\,\mathrm{m}^2$ in the five-interval case. ■

Exercise 1

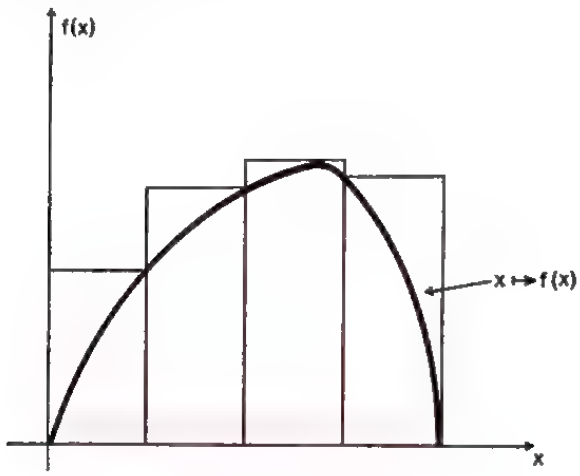
Exercise 1 (5 minutes)

In the two examples discussed in the text, we remarked that the difference between our upper and lower estimates of the required area (using rectangles) was simply the area of the largest rectangle (see pages 8 and 14). Can we say this when trying to find the area under *any* curve by using rectangles, or was there something special about these two examples? ■

Solution 1

Solution 1

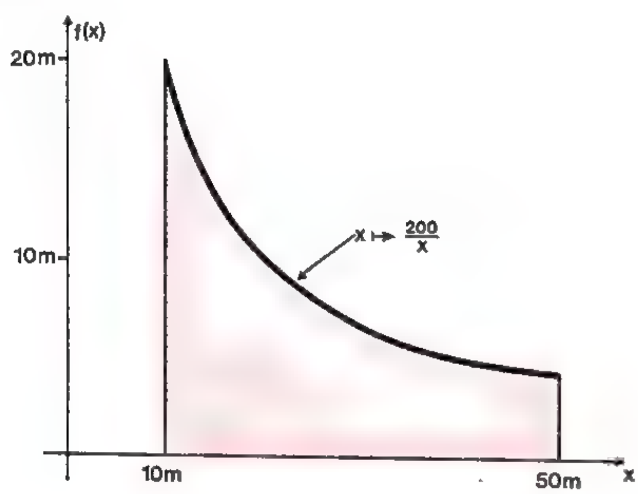
The statement we made applies in the general case *only* if the graph of the function slopes downwards or upwards over the whole region. For example, the statement is not true for the diagram shown below, although we can still use rectangle approximations to find the area. One set of rectangles could be chosen, the sum of whose areas is greater than the required area, and the other set so that the sum of the areas of these rectangles is smaller than the required area. An example of the first set of diagrams is shown in the figure.



■

Exercise 2

Exercise 2
(5 minutes)



The boundary of the cross-section of the roof of a main airport building is the graph of the function

$$f : x \longmapsto \frac{200}{x} \qquad (x \in [10, 50])$$

The supporting walls are vertical and situated at $x = 10$ m and $x = 50$ m respectively.

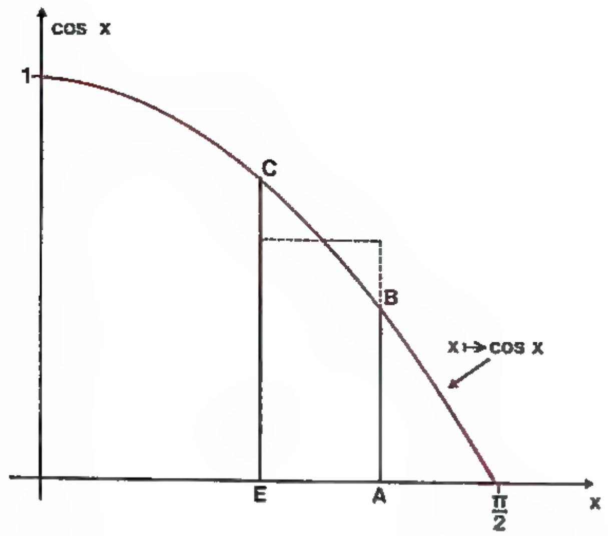
(i) Using four intervals, estimate the cross-sectional area (indicated in the diagram) as accurately as you can by the method we have used in the example, and state the accuracy which you think you have attained.

(ii) How many intervals would you require to attain an estimate with absolute error bound 1 m^2?

■

Exercise 3

Exercise 3
(4 minutes)



Given the function

$$f : x \longmapsto \cos x \qquad \left(x \in \left[0, \frac{\pi}{2}\right]\right)$$

we can find an estimate for the area between its graph and the two axes, by forming rectangles in the following way:

We divide the domain into equal intervals and take the height of the rectangle on each interval to be $\frac{1}{2}$ (the sum of the ordinates* of the end points of the interval). E.g. the height of the rectangle on $EA = \frac{1}{2}(AB + EC)$.

Are the following statements true or false?

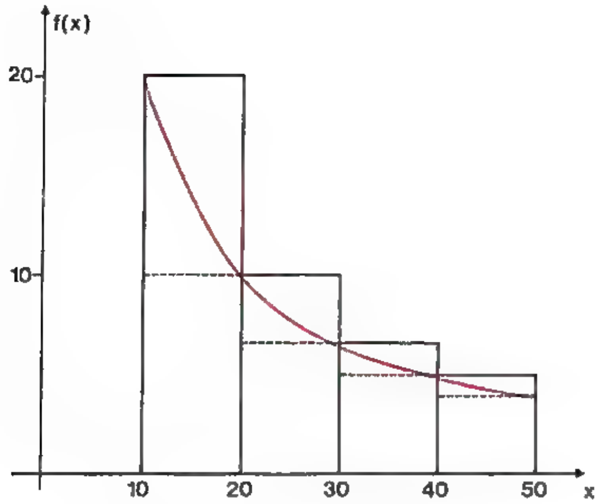
(i) The sum of the areas of the rectangles above is the same as the estimate of the area which we have used so far. That is, the estimated area is the same as that obtained by the previous method: $\frac{1}{2}$ (the sum of the areas of the larger rectangles + the sum of the areas of the smaller rectangles). TRUE/FALSE

(ii) The previous method of estimating the area is better than the method described in this exercise because it gives us an estimate of the magnitude of the error. TRUE/FALSE

■

* The ordinate of a point in the Cartesian plane is the y-co-ordinate of the point.

Solution 2



(i)

x	10	20	30	40	50
$f(x)$	20	10	6.67	5	4

The total area of the larger rectangles $= 416.7\,\text{m}^2$.

(Notice that in this exercise an inherent error, due to the finite decimal representation of $\frac{2}{3}$, has been introduced. However, we can always make this as small as we please.)

The total area of the smaller rectangles $= 256.7\,\text{m}^2$. The estimate for the required area is therefore $\dfrac{(416.7 + 256.7)}{2}\,\text{m}^2 = 336.7\,\text{m}^2$, with estimated absolute error $= \dfrac{160}{2}\,\text{m}^2 = 80\,\text{m}^2$, so in this case the relative error is large.

(ii) The differences in height and area between the largest and smallest rectangles are 16 m and $16 \times$ (interval width) m^2 respectively, so we require $16 \times$ (interval width) $= 1$. Therefore the width of each rectangle must be $\frac{1}{16}$ m. The total number of intervals required is therefore $\dfrac{40}{\frac{1}{16}} = 640$.

We noticed previously in Example 1 that the difference between the sum of the larger rectangles and the sum of the smaller rectangles is a very crude estimate of the error. In fact, the exact area in this case can be calculated and it is $321.9\,\text{m}^2$ (to one place of decimals), which lies well inside the error interval which we obtained in (i). ■

Solution 3

Solution 3

(i) TRUE. When in our original method we take the estimate $\frac{1}{2}$ (the sum of the areas of the larger rectangles + the sum of the areas of the smaller rectangles) this is exactly the same as the sum of $\frac{1}{2}$ (the area of the larger rectangle + the area of the smaller rectangle) taken over all the intervals.

(ii) TRUE. The original method *also* told us the magnitude of the error we might be making. Although this error estimate may be a wild exaggeration, it is better to have some estimate than none at all. ■

9.1.3 Summary

9.1.3 Summary

In this section we have explored the idea of estimating the (intuitive) area of a region bounded by a curve. The methods we have discussed have considerable practical application, to which we return in section 9.4. For the moment, however, we wish to use the ideas of this section to develop a mathematical definition of area under a curve.

9.2 THE DEFINITE INTEGRAL

9.2

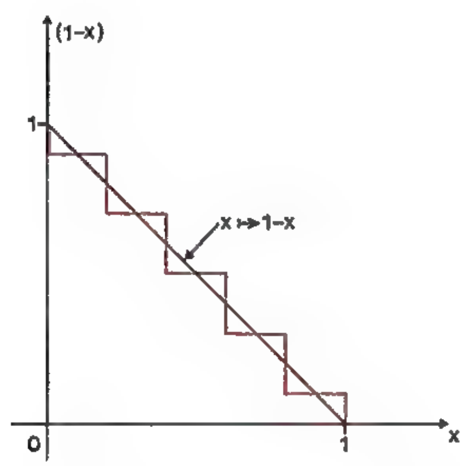
9.2.0 Introduction

9.2.0 Introduction

In section 9.1 we developed a method of approximating to what we intuitively call "area", by considering a number of rectangles, this number being determined by the accuracy required, except in the cases when inherent errors occur. In this section, we shall tie up the intuitive definition of area with a mathematical one based on the ideas you met in section 9.1. We assume that you were satisfied with the statement that, in the absence of inherent errors, increasing the number of rectangles increases the accuracy of the estimate of the area. If so, you were quite justified, but in order to help you to see that we were making an assumption, let us consider the problem of estimating the *length* of a curve.

Exercise 1

Exercise 1
(2 minutes)



We consider a line AB of known length to find out if an estimation procedure, similar to the one we have been considering for areas, gives us a reasonable approximation to the known length.

The graph of the function

$$x \longmapsto 1 - x \qquad (x \in [0, 1])$$

is a straight line, AB.

We approximate the length of this straight line by a "stair-case", the stair treads (intervals) being of equal length. Just as in the area estimation, we can do this in a number of ways, of which we illustrate one.

(i) In the sense of finding the length of a stair carpet, find the total length of the zig-zag line in the figure above.

(ii) Does the number of steps in a stair-case like the one illustrated make any difference to its total length?

(iii) Intuitively does it appear to you that we can get the corners of the stair-case as near to the line AB as we like by taking enough steps?

(iv) Do you think that this implies that we can take enough steps to make the length of the staircase approximate as closely to the length of the straight line as we want?

(v) What then is the length of the line, in the limit, by this procedure?

(vi) Is this the correct length of the line? If not, what is wrong? ■

Discussion

The purpose of this last exercise was not particularly to investigate the idea of length but to show that we must be careful in the limiting procedures we adopt to back up intuitive ideas. Approximating a curve by a "staircase" seems satisfactory when finding the area bounded by the curve, but it is unsatisfactory when trying to find the length of the curve. In both cases we are trying to generalize a concept. Whenever we do this, we must check that our generalized definition gives the answer we expect in the more fundamental case. The "staircase" approach does not give us the "right" answer for the length of a straight line: so there is no point in trying to generalize it. But our method for the calculation of area does satisfy intuition and accords with our concept of area in simple cases. Further, we were able to sandwich the required value between two estimated upper and lower bounds. Thus we can speak of the "accuracy" of our estimate, since we have trapped the numerical value of the area in an error interval which we can make as small as we please, and so we can proceed with confidence.

In section 9.2.1 we shall extract the mathematical definition of the definite integral which we have chosen to use. We say "chosen to use" because there are several ways of defining the definite integral, some of which are more general than others.

In section 9.2.2 we shall show that it is not always necessary to determine definite integrals completely from scratch.

We find that functions such as:

$$x \longmapsto 3x - x^2 + 2 \qquad (x \in R)$$

$$x \longmapsto 2 \sin x - \frac{1}{x} \qquad (x \in R^+)$$

and other similar functions can be expressed in terms of simpler functions, such as:

$$x \longmapsto 1,\ x \longmapsto x,\ x \longmapsto x^2,\ x \longmapsto x^3,\ x \longmapsto \sin x \qquad (x \in R),$$

$$x \longmapsto \frac{1}{x},\ x \longmapsto \frac{1}{x^2} \qquad (x \in R, x \neq 0)$$

(See Part 2 of *Unit 1, Functions*, in which we discussed combinations of functions.) Can we find standard expressions for the definite integrals of simple functions such as these? If so, can we use them to find the definite integrals of more complicated functions, and thus simplify our work? The answer to the first question is "yes". We proceed to find some of these answers in section 9.2.1: in other cases the calculations would be too lengthy at this stage, and so we shall leave them to *Unit 13, Integration II*. The answer to the second question is also "yes" in many cases, such as the examples given above. For combinations of simple functions such as

$$x \longmapsto 3x - x^2 + 2 \qquad (x \in R)$$

we need to develop the rules for combination of the appropriate definite integrals, and this we do in section 9.2.2.

9.2.1 Area and the Definite Integral

9.2.1

Definitions and Notation

Using the experience we have gained in section 9.1, let us try to extract the definition of a definite integral in such a way that it matches up with our intuitive idea of area.

In section 9.1 we took a region and sandwiched it between two sets of rectangles, one containing the region and the other contained in the region. We agreed that the "area" of the region lay between two areas: the sums of the areas of each of the two sets of rectangles.

Suppose that we use this method to obtain upper and lower estimates of the area of the region by considering two sets of rectangles, each with n members.

Notation 1

We denote the sum of the areas of the larger rectangles by A_n and the sum of the areas of the smaller rectangles by a_n. n can be any positive integer, so if we let n take the values $1, 2, 3, \ldots$ successively, we obtain two sequences:

$$A_1, A_2, A_3, \ldots$$

and

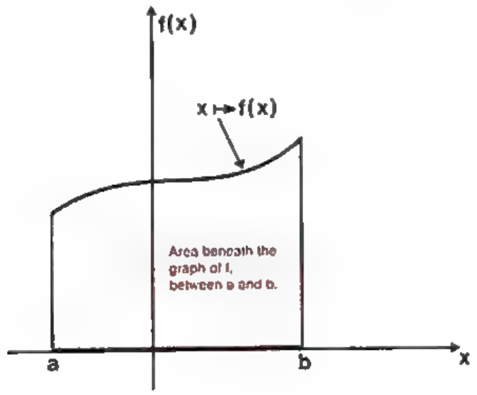
$$a_1, a_2, a_3, \ldots$$

Notation 2

which we denote by $\underset{\sim}{A}$ and $\underset{\sim}{a}$ respectively. (See the notation for infinite sequences introduced in *Unit 7*.)

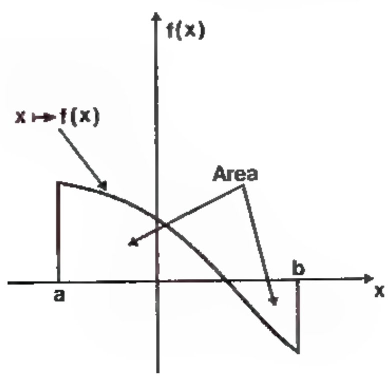
Definition 1

The terms of $\underset{\sim}{A}$ are always greater than the required area, and the terms of $\underset{\sim}{a}$ are always less than the required area, but as n increases, the corresponding terms in the two sequences (intuitively) get closer and closer. So as the number of rectangles gets very large, we can say (intuitively) that the two sequences have the same limit, and we can *define* this limit to be the area of the region. We have to be a bit more specific about area when we think in terms of more general functions. Let f be a function with domain $[a, b]$, such that the graph of f, the lines specified by $x = a$ and $x = b$, and the x-axis form the boundary of a closed region.

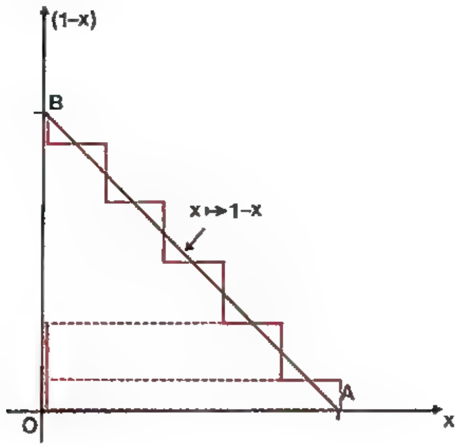


Definition 2

The area of this region is usually called the area beneath the graph of f between a and b even though the graph of f may look like this:



(*continued on page 22*)



(i) Each rising tread is equal to a corresponding length on the $(1 - x)$-axis. These lengths do not overlap, and therefore they add up to the length OB.

Each step length is equal to a corresponding length on the x-axis, and again the lengths do not overlap, but add up to the length OA. So

$$\text{the total length of the zig-zag line} = \text{the total length of the carpet required}$$
$$= 1 + 1 = 2$$

(ii) NO. The total length remains unchanged no matter how many steps are used to cover the line.

(iii) Intuitively YES.

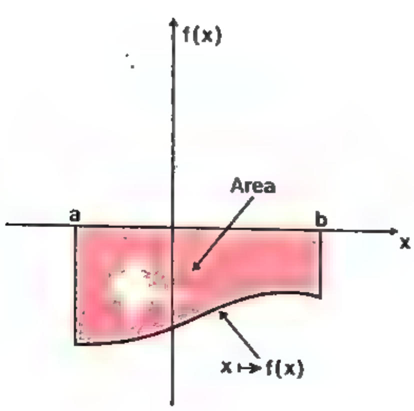
(iv) You have probably answered YES.

(v) Apparently 2 units, using (i) and (ii).

(vi) NO. The correct length is $\sqrt{2}$ units. The stair-carpet, or zig-zag line, is always either horizontal or vertical and is never lying flat on the slope of the hypotenuse of the triangle. So, although the answer to (iii) is correct, it is misleading. ■

(*continued from page 21*)

or this:



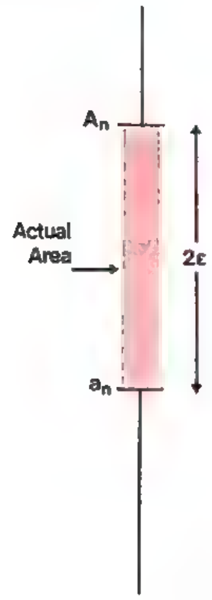
At the end of the example about the Dollan swimming baths (page 15) we found that to achieve an estimate of area with absolute error bound

$\varepsilon\ m^2$ we would require at least n intervals, where n is an integer and

$$n \geqslant \frac{1000}{\varepsilon}$$

That is, we found we could choose n such that the estimate of the area belonged to an error interval of width at most 2ε, where ε could be as small as we pleased. Here we have a similar case.

Expressed the other way round, this means that as n increases the width of the error interval decreases.



For n intervals, the upper and lower bounds of the error interval are determined by A_n and a_n respectively.

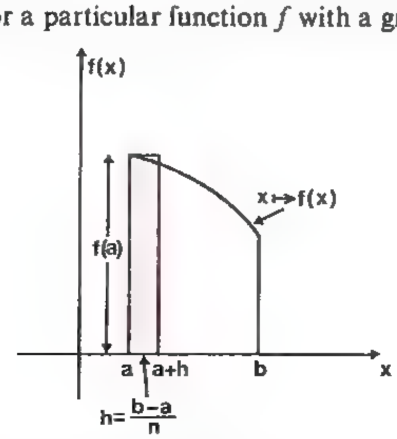
If $\lim \underset{\sim}{A}$ exists, then $\lim \underset{\sim}{a}$ will also exist, and vice versa. The two limits will be the same number: call it A. As n increases, the width* of the error interval will shrink towards zero, and the area of the parabolic section, which is always contained in the error interval, must (intuitively) therefore also be A. Notation 3

Now let us consider the problem for the function f more algebraically, and pick a particular sum (you will see why in a moment):

$$\begin{aligned} S_n &= hf(a) + hf(a+h) + \cdots + hf(a + \{n-1\}h) \\ &= h[f(a) + f(a+h) + \cdots + f(a + \{n-1\}h)] \end{aligned}$$

Definition 3

where $h = \dfrac{b-a}{n}$; that is, we have divided $[a, b]$ into n sub-intervals† each of width h. This expression for S_n defines a sequence $S_1, S_2, S_3, \ldots$ which we denote by $\underset{\sim}{S}$. For a particular function f with a graph as illustrated: Notation 4

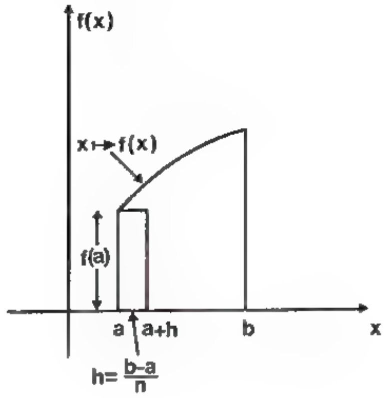


* The width of $[a, b]$ is $b - a$.
† A sub-interval of $[a, b]$ is an interval $[c, d]$ where $a \leqslant c < d \leqslant b$.

we would have

$$S_n = A_n,$$

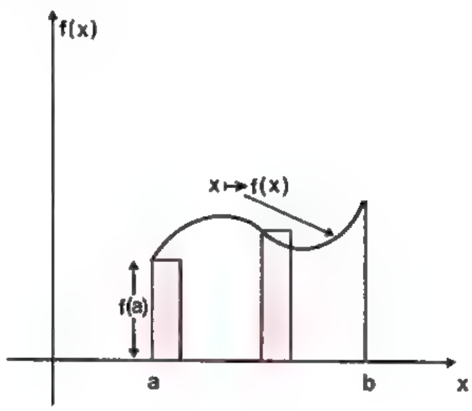
the sum of the areas of the larger rectangles. If f has a graph like this:



then

$$S_n = a_n$$

If f has a graph like this:



then

$$a_n < S_n < A_n$$

The point is this: if the graph of f is a smooth curve, and $f(x)$ is positive* for all values of x in $[a, b]$, then, provided that

$$\lim \underset{\sim}{a} = \lim \underset{\sim}{A} = A$$

it follows that

$$\lim \underset{\sim}{S} = A$$

since S_n will be sandwiched in an error interval which diminishes in size as n gets larger (just like the estimate of the parabolic area in the Dollan Baths example).

Because A, the limit of $\underset{\sim}{S}$, has tremendously important uses in contexts other than that of area, we give it a special name. We call it the definite integral of the function f between a and b (or in $[a, b]$), and a special symbol

Definition 4

$$\int_a^b f$$

Notation 5

* The reason for requiring the images to be positive at this stage will become clear from a consideration of Exercise 1 which follows.

so that

$$\lim \underset{\sim}{S} = \int_a^b f$$

The symbol $\int$, an elongated "s", represents the summation process with the "a" showing where the sum begins and the "b" showing where it finishes. (The "a" and "b" are called the end-points* of the integral).

Definition 5

We can now forget about the rectangles, and concentrate on this particular sequence of sums, $\underset{\sim}{S}$, and its limit. We do this because we are interested not only in area but also in the importance of the definite integral in mathematics and applications. Sometimes the integral will represent an area, but on other occasions it may represent a volume, or an average, or an electric current, or a probability, or power, or distance, or a similar quantity. In each case, the definite integral will have to be interpreted with caution. To avoid some pitfalls we shall stipulate one condition on f which *guarantees* the existence of $\int_a^b f$. Remember that not all sequences converge (i.e. have a limit) and we have defined the definite integral to be the limit of a sequence. We do not want to have to return to the definition repeatedly to check on the existence of integrals. So we ask: Are there any special types of function f for which $\int_a^b f$ always exists? The answer is "Yes", and in fact we are already on solid ground. In our definition on page 24 we have included the phrase "if the graph of f is a smooth curve"; that is, the graph of f must have no "gaps" in $[a, b]$. This is the same as saying that "the function f must be continuous in $[a, b]$" (see *Unit 7, Sequences and Limits I*). It can be shown (by a more rigorous treatment than ours) that *if* f is continuous in $[a, b]$, then the definite integral $\int_a^b f$ *automatically* exists (i.e. its existence is *guaranteed*). We ask you to accept this fact. (It means that the words "provided that" before "$\lim \underset{\sim}{a} = \lim \underset{\sim}{A} = A$" on page 24 are redundant and may go home.) Hereafter, in this text, we shall assume that the definite integral exists (i.e. the sequence $\underset{\sim}{S}$ converges) in all our discussions.

Summary

Summary

Given an interval $[a, b]$ which is divided into n equal sub-intervals of width

$$h = \frac{b-a}{n}$$

and a function f which is continuous in $[a, b]$, we define the definite integral of f in $[a, b]$ as the limit of the sequence $\underset{\sim}{S}$, where

$$S_n = h[f(a) + f(a+h) + \cdots + f(a + \{n-1\}h)]$$

We write

$$\lim \underset{\sim}{S} = \int_a^b f$$

Exercise 1

Exercise 1
(3 minutes)

Is "the definite integral of f between a and b" synonymous with "the area beneath the graph of f between a and b"?

HINT: Consider the two definitions for the diagrams on page 21. ■

* Some authors use the word *limits*.

Solution 1

Solution 1

Your solution should contain the following points:

(i) Even if the curve is wholly above the x-axis, if it goes up *and* down then the two sequences $\underset{\sim}{A}$ and $\underset{\sim}{S}$ are not the same, but, intuitively, the terms of the sequence $\underset{\sim}{S}$ lie between the terms of the two sequences $\underset{\sim}{a}$ and $\underset{\sim}{A}$. If these latter two sequences have the same limit, then so also has $\underset{\sim}{S}$ and we may therefore give an intuitive YES.

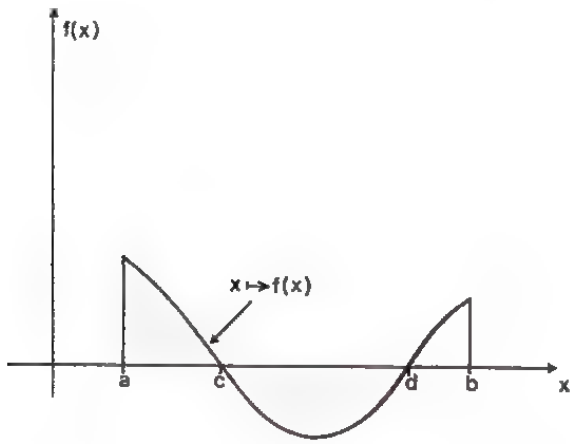
(ii) If some of the curve lies below the x-axis, then we have negative terms in $\underset{\sim}{S}$, hence lim $\underset{\sim}{S}$ is definitely *not* the area. So the definite integral and the area may both exist and be different. ■

Discussion

We may, therefore, conclude that as long as we are careful to ascertain exactly where the curve representing f lies, we can use the definite integral to find the area under the graph of f between a and b. If $f(x)$ is not positive for all $x \in [c, d]$, where $[c, d]$ is a sub-interval of $[a, b]$ (see diagram), then we find the definite integral of f between

$$[a, c], \qquad [c, d], \qquad [d, b]$$

separately, and adjust the sign of the definite integral of f between c and d before adding the three results to find the total area.



The following exercise is designed to give you practice in the ideas introduced in this section and to check that they are in accordance with intuition.

Exercise 2

Exercise 2
(5 minutes)

Find the definite integral of the function f in $[a, b]$, where

$$f: x \longmapsto 1 \qquad (x \in [a, b],\ b > a > 0)$$

and check whether it does give the area beneath the graph of f. ■

Exercise 3

Exercise 3
(3 minutes)

Find the definite integral of the function f in $[a, b]$, where

$$f: x \longmapsto x \qquad (x \in [a, b],\ b > a > 0)$$

and check whether it does give the area beneath the graph of f between a and b.

HINT: In the solution you will need the fact demonstrated in *Unit 4, Finite Differences*, page 47, that the sum of the first n natural numbers is

$$S_1(n) = \frac{n(n+1)}{2}$$ ■

Just a word on notation. We have written the definite integral of a function f in $[a, b]$ as Notation 1

$$\int_a^b f$$

In particular cases when f is known, for example:

$$f: x \longmapsto x \qquad (x \in [a, b])$$

we write

$$\int_a^b x \longmapsto x, \quad \text{or} \quad \int_a^b (x \longmapsto x)$$

and omit the domain of f because the part of the domain we are interested in is clear from the "a" and "b" in the notation.

In many text books you will find the notation

$$\int_a^b f(x)\, dx$$

or, in particular,

$$\int_a^b x\, dx$$

This is a notation which has been used for many years. If you are interested in its origins, you will find it in a book on the history of the subject. The reasons for our choice of notation are discussed in the introduction at the beginning of this text.

The definite of definite integral we have used is not the most general one. We have, for example, chosen the particular case where we divide $[a, b]$ into equal intervals. The definite integral of f in $[a, b]$ can be satisfactorily defined in a more general way, omitting the conditions: "f is continuous in $[a, b]$" and "the sub-intervals of $[a, b]$ are of equal width". On the whole we have relied more on intuition rather than mathematical rigour; we leave the rigorous definition of the definite integral to a later course.

In the last two exercises you found that Main Text

$$\int_a^b (x \longmapsto 1) = b - a \qquad \text{and} \qquad \int_a^b (x \longmapsto x) = \frac{b^2 - a^2}{2}$$

$$b > a > 0$$

By a similar method to that used in the exercises, it can be shown that

$$\int_a^b (x \longmapsto x^2) = \frac{b^3 - a^3}{3}$$

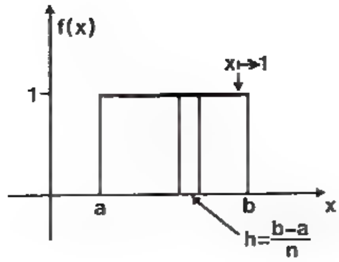
To derive this result, one needs the formula for the sum of the squares of the first n natural numbers

$$S_2(n) = \frac{n(n + 1)(2n + 1)}{6}$$

which can be found in *Unit 4, Finite Differences*, page 48. (*continued on page 29*)

Solution 2

Solution 2



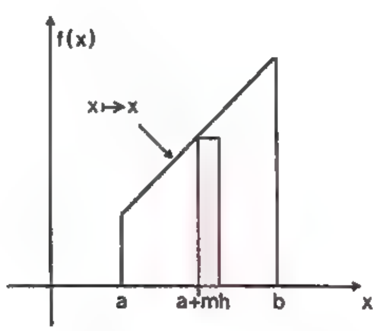
The interval is divided into n sub-intervals of width h, where $nh = b - a$.
Then

$$S_n = h[f(a) + f(a+h) + \cdots + f(a + \{n-1\}h)]$$
$$= \underbrace{h[1 + 1 + \cdots + 1]}_{(n \text{ terms})} = nh = b - a$$

The definite integral between a and b is the limit of the sequence $\underline{S}$, that is $(b - a)$.
This is clearly the area beneath the graph between a and b. ■

Solution 3

Solution 3



The interval is divided into n sub-intervals of width h, where $nh = b - a$.
We have $f(x) = x$, so that

$$f(a) = a, f(a+h) = a + h, \ldots, f(a + \{n-1\}h) = a + \{n-1\}h$$

Therefore

$$S_n = h[f(a) + f(a+h) + f(a+2h) + \cdots + f(a + \{n-1\}h)]$$
$$= h[a + (a+h) + (a+2h) + \cdots + (a + \{n-1\}h)]$$
$$= nha + h^2[1 + 2 + 3 + \cdots + \{n-1\}]$$

(collecting together like terms)

$$= a(b-a) + \frac{(b-a)^2}{n^2}[1 + 2 + 3 + \cdots + \{n-1\}]$$

(substituting for h)

$$= a(b-a) + \frac{(b-a)^2}{n^2} S_1(n-1)$$

$$= a(b-a) + \frac{(b-a)^2}{n^2}\,\frac{(n-1)n}{2} \quad \text{(substituting for } S_1(n-1))$$

$$= a(b-a) + \frac{(b-a)^2}{2}\left(1 - \frac{1}{n}\right)$$

The limit of this sequence for n large is

$$a(b-a)+\frac{(b-a)^2}{2}=\frac{(b^2-a^2)}{2} \qquad \left(\text{since } \lim_{n \text{ large}}\left(1-\frac{1}{n}\right)=1\right)$$

So the definite integral of f in $[a, b]$ is $\frac{1}{2}(b^2-a^2)$.

By regarding the area beneath the graph of f between a and b as the difference between the areas of the larger triangle, whose base is the x-axis from the origin to b, and the smaller triangle, whose base is the x-axis from the origin to a, we see that this area is

$$\tfrac{1}{2}b^2-\tfrac{1}{2}a^2$$

which is the same as the definite integral of f in $[a, b]$. ■

(*continued from page 27*)

Exercise 4

Exercise 4
(2 minutes)

(i) From the results so far obtained, an intuitive guess at the value of

$$\int_a^b (x \longmapsto x^m), \qquad b>a>0, \quad m \in Z^+$$

might be

(a) $\dfrac{b^{m+1}-a^{m+1}}{m}$

(b) $\dfrac{b^m-a^{m+1}}{m}$

(c) $\dfrac{b^{m+1}-a^{m+1}}{m+1}$

(d) $\dfrac{b^m-a^m}{m}$

Which do you think is correct?

(ii) To obtain the above value would you need to know the sum of the

(a) $(m-1)$th

(b) mth

(c) $(m+1)$th

powers of the first n natural numbers? ■

Exercise 5

Exercise 5
(2 minutes)

Calculate, using the general result for the definite integral of $x \longmapsto x^m$, obtained as the correct answer to the previous exercise, the values of

(i) $\displaystyle\int_0^4 (x \longmapsto x^3)$

(ii) $\displaystyle\int_{-2}^2 (x \longmapsto x^4)$

(iii) $\displaystyle\int_0^1 (x \longmapsto x^{43})$ ■

Solution 4

Solution 4

(i) (c)
(ii) (b) ■

Solution 5

Solution 5

(i) 64
(ii) $\frac{64}{5}$
(iii) $\frac{1}{44}$ ■

Discussion

You have probably recognized that the calculation of

$$\int_a^b (x \longmapsto x^m) \qquad (m \in Z^+)$$

from first principles becomes more cumbersome as m increases in magnitude. That is why we leave it until we get some tools in *Unit 13, Integration II*, which can handle it very simply. The formula we obtain if we did the calculation is, as already noted,

$$\int_a^b (x \longmapsto x^m) = \frac{b^{m+1} - a^{m+1}}{m+1} \qquad (m \in Z^+)$$

In fact, although we cannot show it at this stage, this formula is true for $(m \in R, m \neq -1)$, and you may assume this for any subsequent exercises in this unit.

Table of Definite Integrals of
Simple Polynomial Functions

f	$\int_a^b f$
$x \longmapsto 1$	$b - a$
$x \longmapsto x$	$\frac{b^2 - a^2}{2}$
$x \longmapsto x^2$	$\frac{b^3 - a^3}{3}$
$x \longmapsto x^m \ (m \in R, m \neq -1)$	$\frac{b^{m+1} - a^{m+1}}{m+1}$

The first two above were derived in the text; we indicated how to prove the third, and you may use the general result although we have not proved it.

9.2.2 The Definite Integral for Combinations of Functions

9.2.2

Main Text

In order to extend the class of functions for which we can calculate definite integrals to include, for example,

$$\int_1^2 x \longmapsto (3x^3 + 5x)$$

we need two theorems (unless, of course, we resort to summing series!). We derive one of them in the following example and the other in the exercise below.

Example 1

Example 1

Show that

$$\int_a^b cf = c\int_a^b f$$

where f is a function, c is any number and cf is the function

$$x \longmapsto cf(x)$$ ■

Solution of Example 1

This is not such a rigorous proof as that which appears in B. Hunt, *Calculus and Linear Algebra* (W. H. Freeman 1967), page 273 (see Bibliography), but the essence of the argument is here.

To define the definite integral of f between a and b we used the limit of the sequence $\underset{\sim}{S}$, where

$$S_n = h[f(a) + f(a + h) + \cdots + f(a + \{n - 1\}h)]$$

Similarly, to define the definite integral of (cf) between a and b we would use the limit of the sequence $\underset{\sim}{T}$, where

$$\begin{aligned} T_n &= h[cf(a) + cf(a + h) + \cdots + cf(a + \{n - 1\}h)] \\ &= cS_n \end{aligned}$$

It follows from the result on page 33 of *Unit 7, Sequences and Limit I*, that

$$\lim \underset{\sim}{T} = c \lim \underset{\sim}{S}$$

and hence

$$\int_a^b cf = c\int_a^b f$$ ■

Exercise 1

Exercise 1
(3 minutes)

Convince yourself, by using the definition of the definite integral, that if $f(x)$ and $g(x)$ are positive for all values of x in $[a, b]$, then

$$\int_a^b (f + g) = \int_a^b f + \int_a^b g$$ ■

In section 7.3.1 of *Unit 7, Sequences and Limits I*, page 34, we intuitively derived the result

$$\lim (\underset{\sim}{S} + \underset{\sim}{T}) = \lim \underset{\sim}{S} + \lim \underset{\sim}{T}$$

(A more rigorous proof was given in an appendix.) In that section of *Unit 7* we interpreted this result in terms of morphisms, and we may similarly interpret our present result.

(continued on page 32)

Solution 1

Write

$$S_n = h[f(a) + f(a + h) + \cdots + f(a + \{n - 1\}h)]$$
$$T_n = h[g(a) + g(a + h) + \cdots + g(a + \{n - 1\}h)]$$

and take the appropriate limits. ■

(*continued from page 31*)

Let F be the set of all functions f for which it is possible to evaluate

$$\int_a^b f$$

If a and b are fixed real numbers, then the function I, defined by

$$I: f \longmapsto \int_a^b f \qquad (f \in F)$$

maps the elements of F to the set of real numbers.

The result of this exercise then becomes

$$I(f + g) = I(f) + I(g)$$

where the $+$ on the left represents the addition of functions, and the $+$ on the right the addition of real numbers. We can express this result by saying that I is a morphism of $(F, +)$ to $(R, +)$.

The appropriate commutative diagram is

$$\begin{tikzcd}
(f, g) \arrow[r, "+"] \arrow[d, dashed, "I"'] & f + g \arrow[d, "I"] \\
(I(f), I(g)) \arrow[r, "+"] & I(f) + I(g) = I(f + g)
\end{tikzcd}$$

Noticing that, for instance,

$$\int_0^1 x \longmapsto 1 = 1$$

and

$$\int_0^1 x \longmapsto 2x = 1,$$

we see that I is many-one, and hence it is a homomorphism. In fact, we could have approached this section from this point of view, first noticing the function I and then investigating whether it is compatible with any of the known binary operations on F, for example, $+$, $\times$ and $\circ$ (the composition of functions). The answers for $\times$ and $\circ$ is that it is not compatible.

Example 2

Example 2

Calculate

$$\int_1^3 x \longmapsto (3x^3 + 5x)$$

■

Solution of Example 2

$$\int_1^3 (x \longmapsto (3x^3 + 5x))$$

$$= \int_1^3 (x \longmapsto 3x^3) + \int_1^3 (x \longmapsto 5x), \text{ from rule in Exercise 1}$$

$$= 3\int_1^3 (x \longmapsto x^3) + 5\int_1^3 (x \longmapsto x), \text{ from rule in Example 1}$$

$$= 3\left(\frac{3^4 - 1^4}{4}\right) + 5\left(\frac{3^2 - 1^2}{2}\right)$$

$$= 80$$

■

Exercise 2

Exercise 2
(3 minutes)

Calculate

(i) $\int_2^4 x \longmapsto (2x^2 + 7x - 3)$

(ii) $\int_{-2}^2 x \longmapsto (4 - x^2)$

(iii) $\int_2^3 x \longmapsto (x^2 - x)$

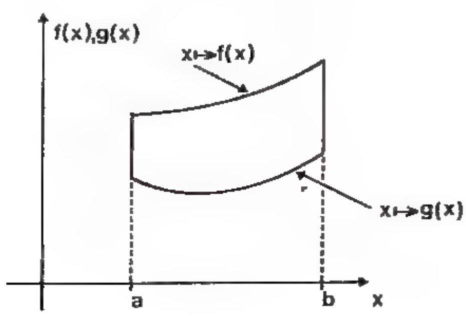
(iv) $\int_0^3 x \longmapsto (x - 1)(x - 2)$

(v) What is the area beneath the graph of f in (iv)? ■

Main Text

It is frequently possible to find areas of more complex regions with very little extra effort.

Suppose we wish to find areas which are bounded by curves and lines specified by equations of the form $x = c$, $c \in R$, but are not bounded by the x-axis; for example, the area shaded in red in the diagram.



(continued on page 35)

Solution 2

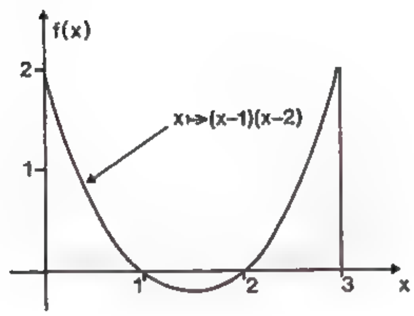
(i) $\int_2^4 (x \longmapsto (2x^2 + 7x - 3))$

$$= 2\int_2^4 (x \longmapsto x^2) + 7\int_2^4 (x \longmapsto x) - 3\int_2^4 (x \longmapsto 1)$$

$$= 2\left(\frac{4^3 - 2^3}{3}\right) + 7\left(\frac{4^2 - 2^2}{2}\right) - 3(4 - 2)$$

$$= \frac{220}{3}$$

(ii) $\frac{32}{3}$
(iii) $\frac{23}{6}$
(iv) $\frac{3}{2}$
(v)



The graph of the function is shown in the diagram, and because part of the graph of this function lies below the x-axis, the definite integral evaluated in (iv) does not give the required area. So we divide the definite integral into 3 separate limiting sequences. In the first and third sequence all the terms are positive whilst in the second they are negative. The appropriate definite integrals are

$$\int_0^1 (x \longmapsto (x^2 - 3x + 2)) = \tfrac{5}{6}$$

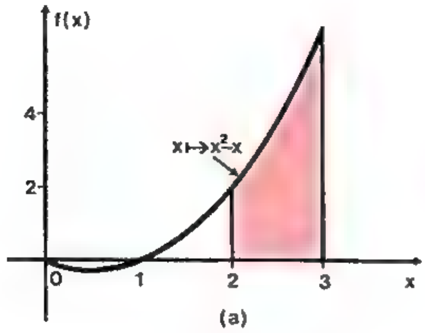
$$\int_1^2 (x \longmapsto (x^2 - 3x + 2)) = -\tfrac{1}{6}$$

$$\int_2^3 x \longmapsto (x^2 - 3x + 2) = \tfrac{5}{6}$$

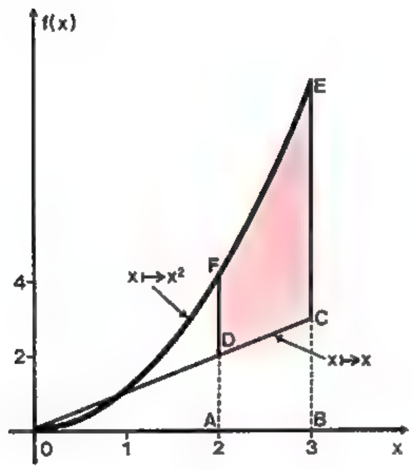
The *area* is the sum of the *magnitudes* of these three definite integrals $= \frac{5}{6} + \frac{1}{6} + \frac{5}{6} = \frac{11}{6}$. ■

Rather than try to form rectangles to determine the numerical value for this area, we can consider it as the difference of two areas bounded by the x-axis. Look more closely at part (iii) of the last exercise from a graphical point of view.

(*continued from page 33*)



(a)



(b)

Since the images of x under the functions $x \longmapsto (x^2 - x)$, $x \longmapsto x^2$, $x \longmapsto x$ are not negative for x in $[2, 3]$, the definite integral

$$\int_2^3 x \longmapsto (x^2 - x)$$

represents the area shaded in red in diagram (a), and the definite integrals

$$\int_2^3 (x \longmapsto x^2), \qquad \int_2^3 (x \longmapsto x)$$

represent the areas $ABEF$ and $ABCD$ respectively in diagram (b).
The shaded areas in (a) and (b) may not *look* equal, but since

$$\int_2^3 (x \longmapsto (x^2 - x)) = \int_2^3 (x \longmapsto x^2) - \int_2^3 (x \longmapsto x)$$

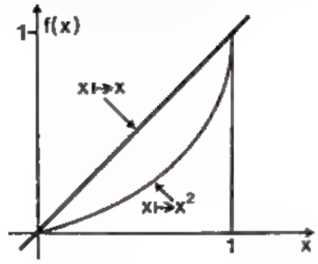
these areas are the same.

Exercise 3

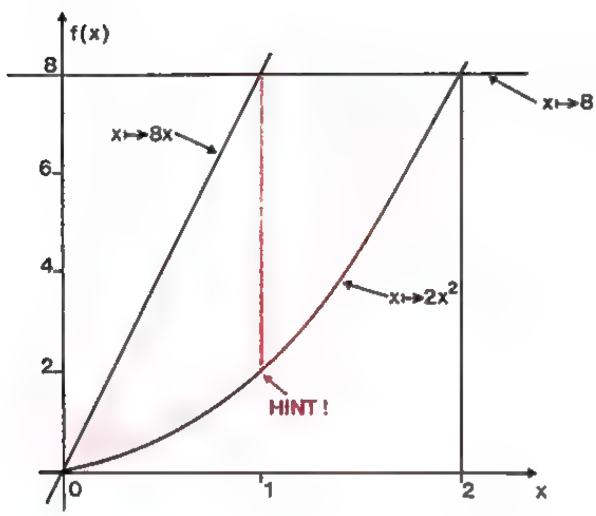
Exercise 3
(5 minutes)

Write down the definite integral, or sum of definite integrals, which represents the areas on the following graphs, and hence evaluate them.

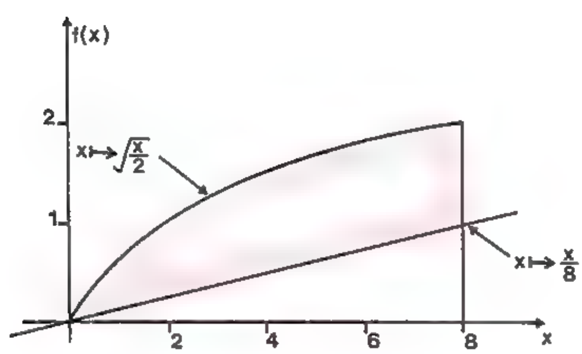
(i)



(ii)



(iii)



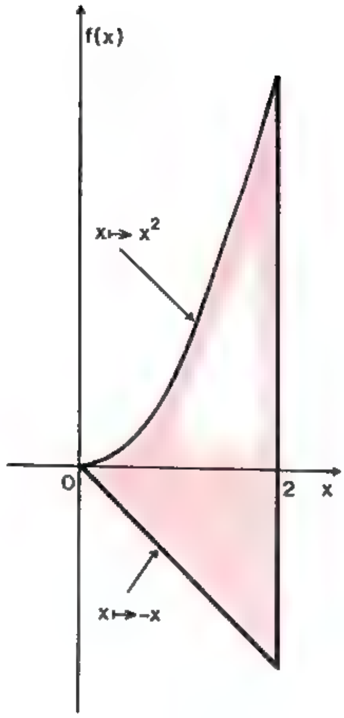
For this one, you will need the result

$$\int_a^b x \longmapsto x^{1/2} = \tfrac{2}{3}(b^{3/2} - a^{3/2})$$

(iv) Is it a coincidence that (ii) and (iii) have the same answer? ■

Exercise 4

Exercise 4
(3 minutes)

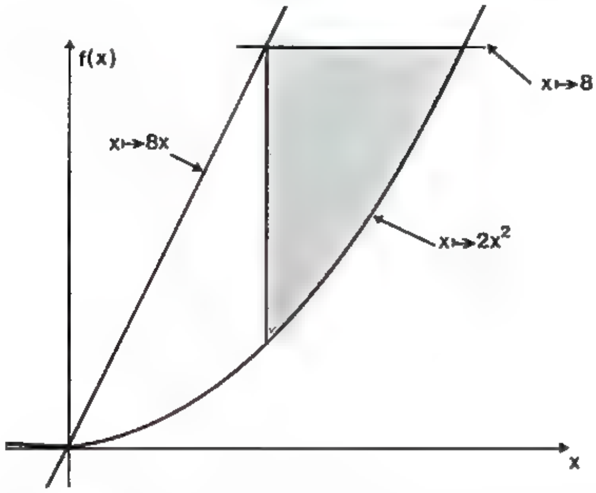


If the shaded area represents a piece of metal 1 cm thick, and x and its images are measured in centimetres, calculate the volume of material in such a metal plate. ■

Solution 3 Solution 3

(i) $\int_0^1 x \longmapsto (x - x^2) = \frac{1}{6}$

(ii) $\int_0^1 x \longmapsto (8x - 2x^2) + \int_1^2 x \longmapsto (8 - 2x^2) = \frac{20}{3}$

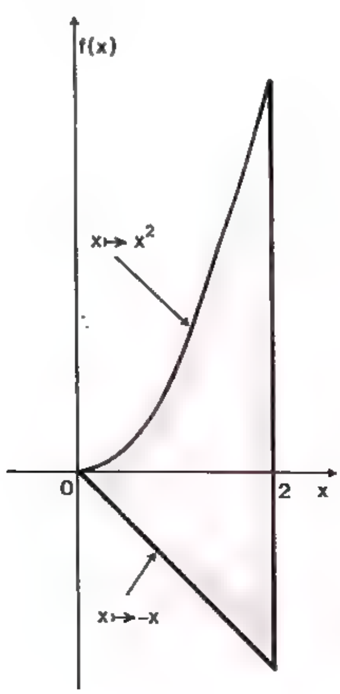


The first definite integral represents the area shaded in red, the second the area shaded in black. First find the appropriate intersections of the graphs, then write down the integrals to give the result above.

(iii) $\int_0^8 x \longmapsto \left(\frac{1}{\sqrt{2}} \times x^{1/2} - \frac{x}{8}\right) = \frac{20}{3}$

(iv) NO. In (iii) we have drawn the graphs of the inverses of the functions used in (ii). Effectively we could have used the same diagram as in (ii) and interchanged the axes. We often face the choice of forming the definite integral "the other way", that is, using the inverse functions. The simplicity gained in the representation of the area can, however, be lost in the awkwardness of the new function, as we see in this example. ■

Solution 4 Solution 4



If the areas above and below the x-axis are calculated separately by definite integrals, we have

$$\int_0^2 (x \longmapsto x^2) + \text{the magnitude of} \left[\int_0^2 (x \longmapsto -x)\right]$$

$$= \tfrac{8}{3} + 2 = \tfrac{14}{3}$$

so the volume is $\frac{14}{3}$ cm^3. We can also represent the whole area by the one definite integral

$$\int_0^2 x \longmapsto (x^2 - (-x)) = \int_0^2 x \longmapsto (x^2 + x) = \tfrac{14}{3},$$

and again we find that the volume is $\frac{14}{3}$ cm^3. ■

9.3 SOME APPLICATIONS OF THE DEFINITE INTEGRAL

9.3

9.3.0 Introduction

9.3.0

Introduction

We have defined the definite integral of f in $[a, b]$ as the limit of the sequence

$$S_1, S_2, S_3, \ldots, S_n, \ldots$$

where

$$S_n = h[f(a) + f(a + h) + \cdots + f(a + \{n - 1\}h)]$$

We have also shown that provided the images are positive, this integral will give us the area beneath the graph of f between a and b. In this section we intend to look at other physical situations which, when we analyze them, produce sums of terms which we can identify with a definite integral. For example, the terms in the sequence $S_1, S_2, S_3, \ldots, S_n, \ldots$, which were used to represent the sums of the areas of rectangles, can be used to represent the sums of volumes of standard shapes. We look at this in more detail in the next section.

9.3.1 Volume of a Solid of Revolution

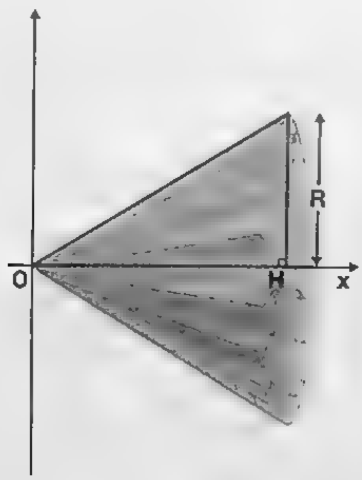
9.3.1

Main Text

Consider a region bounded by the graph of a function f (whose images are positive in $[a, b]$), the lines specified by $x = a$ and $x = b$, and the interval of the x-axis between a and b. Focus attention on the boundary of such a region, and imagine the boundary being rotated round the x-axis; so that it generates the bounding surface of a solid. We usually refer to the rotation of the graph of f in $[a, b]$ (leaving the rotation of the area to be understood.) The volume of such a solid is called a volume of revolution, and the calculation of these volumes is simply an application of the definite integral.

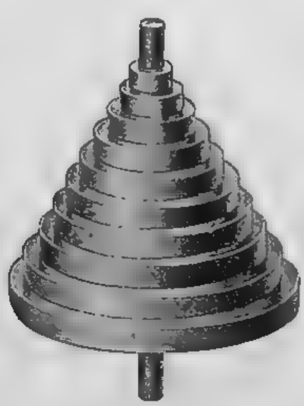
Definition 1

As an illustration, we shall investigate the problem of finding the volume of a cone generated by rotating the hypotenuse of a right-angled triangle around the x-axis.



You may know that the volume, V, is given by the formula $V = \frac{1}{3}\pi R^2 H$ where H is the height of the cone and R is the radius of the base. We shall derive this formula.

To correspond to our elementary rectangles (used in the calculation of area) we choose discs of thickness h. They are similar to the discs on the child's toy shown in the diagram.

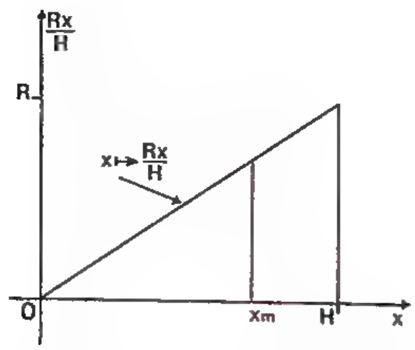


If there are n discs we have

$$h = \frac{H}{n},$$

and the edge of the cone will be the graph of the function

$$x \longmapsto \frac{Rx}{H} \qquad (x \in [0, H])$$



The radius of each of the n discs will be given by an ordinate. Let the n ordinates be at $x_0, x_1, \ldots, x_{n-1}$, then

$$x_m = mh \qquad (m = 0, 1, 2, \ldots, n-1)$$

and the corresponding ordinate is the image of x_m, so that

$$\frac{Rx_m}{H} = \frac{Rmh}{H}$$

and the volume of the elementary disc is therefore

$$\pi\left(\frac{Rmh}{H}\right)^2 h$$

The total volume of these n discs is

$$V_n = 0 + \frac{\pi R^2 h^3}{H^2} \times 1^2 + \cdots + \frac{\pi R^2 h^3}{H^2} \times m^2 + \cdots$$
$$+ \frac{\pi R^2 h^3}{H^2} \times (n-1)^2$$
$$= h\left[0 + \frac{\pi R^2}{H^2} h^2 + \cdots + \frac{\pi R^2}{H^2}(mh)^2 + \cdots + \frac{\pi R^2}{H^2}\{(n-1)h\}^2\right]$$

We write the expression for V_n this way because we are going to compare it with the sum whose limit we know to be a definite integral, rather than go through the lengthy algebraic process of finding the sum again. Thus, comparing V_n with S_n, given on page 23:

$$S_n = h[f(a) + f(a+h) + \cdots + f(a+mh) + \cdots$$
$$+ f(a + \{n-1\}h)]$$

it is fairly easy to see that the appropriate function is

$$f : x \longmapsto \frac{\pi R^2}{H^2} x^2$$

and that the end-points of integration are $a = 0$ and $b = nh + a = nh = H$. Thus

$$\lim \underset{\sim}{V} = \lim \underset{\sim}{S} = \int_0^H x \longmapsto \frac{\pi R^2 x^2}{H^2}$$
$$= \tfrac{1}{3}\pi R^2 H$$

Old hands at the game would abbreviate the above comparison even further. They would jump straight from the volume of the elementary disc in the mth position:

$$\pi\left(\frac{Rx_m}{H}\right)^2 h$$

to the appropriate function and then to the integral, by dropping the "h" and the subscript "m". In fact, the justification for doing this is outlined by the fuller argument above.

We can generalize the definite integral for finding volumes of revolution. If the volume of revolution is formed by the graph of f rotating about the x-axis between $x = a$ and $x = b$, then the volume generated is

$$\int_a^b x \longmapsto \pi\{f(x)\}^2 \quad \text{or} \quad \pi \int_a^b x \longmapsto \{f(x)\}^2$$

(We could also use $\pi \int_a^b f^2$ but f^2, which here stands for $f \times f$, could be confused with $f \circ f$.)

In the alternative notation we would have

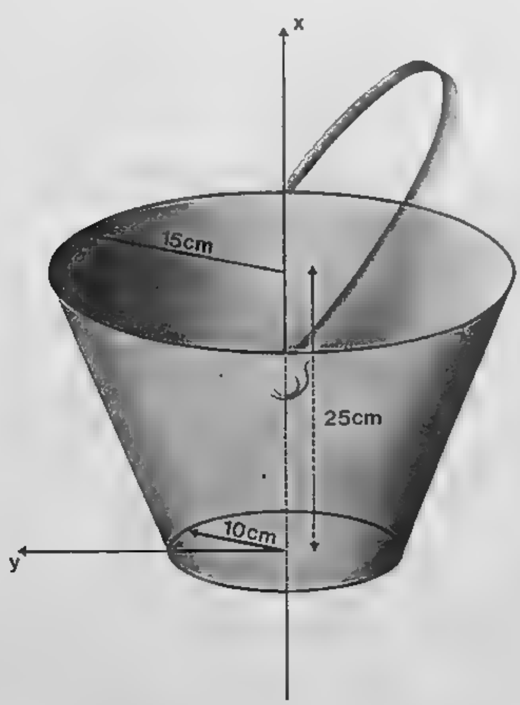
$$\int_a^b \pi\{f(x)\}^2\,dx \quad \text{or} \quad \pi \int_a^b y^2\,dx$$

where $y = f(x)$.

There are two similar exercises to follow in which you are asked to find volumes of solids of revolution. If you are satisfied that you can cope with this type of exercise already, confirm this by tackling just one of the two and then moving on to section 9.3.2.

Exercise 1

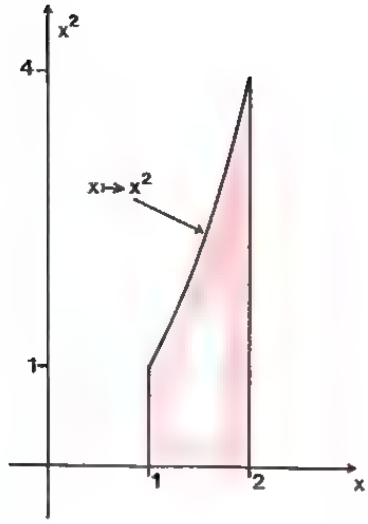
Exercise 1
(3 minutes)



Find the volume of a bucket of circular cross-section with the dimensions shown. ■

Exercise 2

Exercise 2
(2 minutes)



Find the volume within the surface formed by rotating the graph of

$$x \longmapsto x^2 \qquad (x \in [1, 2])$$

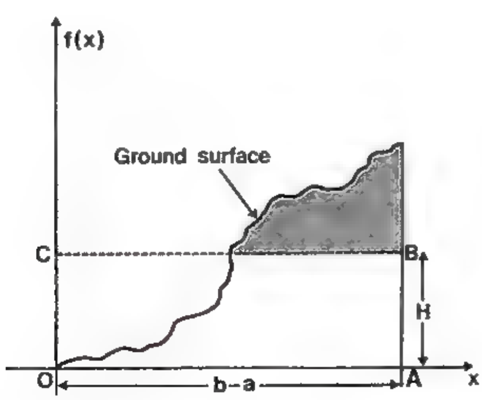
about the x-axis. ■

9.3.2 Averages

9.3.2

Main Text

In section 9.1 we looked at an example of an excavation problem, and determined the volume of earth to be removed. We did this by determining a cross-sectional area using given measurements. In actual practice, when viewing a site for building operations, a builder will often mentally "add a bit here" and "take a bit off there" and take an average depth to estimate the cross-sectional area. For example, faced with the site illustrated below a builder might assume that the area of the red region was the same as the area of the black region, and that the cross-sectional area required was the same as the area of the rectangle $OABC$.

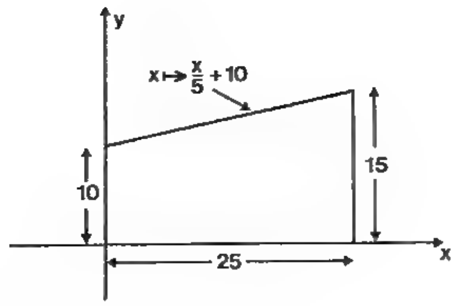


He would probably call the height of this rectangle "the *average* height" of the site. If we suppose that the curve is the graph of a function f between a and b relative to an appropriate axis, then the cross-sectional area is, of course,

$$\int_a^b f$$

(*continued on page 45*)

Solution 9.3.1.1 **Solution 9.3.1.1**



The equation of the bounding curve is given by

$$\frac{y - 10}{15 - 10} = \frac{x - 0}{25 - 0}$$

$$y = \frac{x}{5} + 10$$

[See Refresher Booklet 3 if you cannot follow this.] Therefore the volume expressed as

$$\pi \int_0^{25} x \longmapsto \left(\frac{x}{5} + 10\right)^2$$

$$= \pi \int_0^{25} x \longmapsto \left(\frac{x^2}{25} + 4x + 100\right)$$

$$= \pi \left\{\frac{1}{25}\frac{(25^3 - 0^3)}{3} + 4\frac{(25^2 - 0^2)}{2} + 100(25 - 0)\right\}$$

$$= \pi \left\{\frac{19 \times 625}{3}\right\}$$

$$= 12\,440 \text{ cm}^3$$

$= 12.44$ litres (to 4 significant figures) ■

Solution 9.3.1.2 **Solution 9.3.1.2**

$$\text{Volume} = \pi \int_1^2 x \longmapsto (x^2)^2$$

$$= \frac{\pi(2^5 - 1^5)}{5}$$

$$= \frac{\pi 31}{5}$$ ■

(*continued from page 43*)

The base of our rectangle is then of length $b - a$, and if its height is H, we have

$$\text{average of } f(x) \text{ over } [a, b] = H = \frac{1}{b-a}\int_a^b f$$

Definition 1

which we can abbreviate to:

$$\text{average} = \frac{1}{b-a}\int_a^b f$$

Exercise 1

Exercise 1
(2 minutes)

Find the average of $f(x)$ over $[0, 4]$, where f is the function:

$$x \longmapsto x^2 \qquad (x \in [0, 4])$$

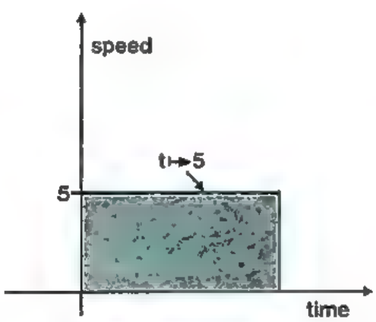
■

9.3.3 Velocity and Distance

9.3.3

Main Text

A hiker goes on a walk lasting four hours, excluding rests. From past experience he reckons that he walks at an average speed of 5 km/h. You conclude that his walk was approximately 20 km long. This was all straightforward since



$$\text{speed} \times \text{time} = \text{distance}$$

you assumed that the speed was constant, and so the calculation was simple. By plotting speed against time on a graph, we see that the area beneath the graph (the area of a rectangle in this case) represents the distance covered.

Similarly, if the speed is not constant, the distance covered is also represented by the area under an appropriate graph. We can see that this is so by dividing the time interval into n equal sub-intervals and then assuming that the speed is constant over each of these time-sub-intervals. The distance travelled during each sub-interval is then represented by the area of a rectangle, and the total distance covered is the sum of the areas of the n rectangles; this will be a close approximation to the area under the graph when n is large.

Here we prefer to use "velocity" rather than "speed" because velocity means speed in a known direction. When we consider motion in a directed straight line, we take velocity to be positive if its direction is the same as that of the line, and negative if it has the opposite direction.

The calculation of the distance travelled is not so easy as it was for the case of the hiker because the required area is given by a definite integral. We must be careful of the physical interpretation of "distance" in relation

(*continued on page 46*)

Solution 9.3.2.1 **Solution 9.3.2.1**

$$\text{Average} = \frac{1}{4-0}\int_0^4 x \longmapsto x^2 = \frac{1}{4}\left(\frac{4^3 - 0^3}{3}\right) = 5\tfrac{1}{3}$$ ■

(continued from page 45)

to the definite integral when the value of the velocity function becomes negative. In any particular context does it mean the *total* distance travelled or the distance that the object is from its starting point? This difference of meaning is illustrated in the following example.

Exercise 1 **Exercise 1** (3 minutes)

A ball, thrown vertically upwards with an initial velocity of 20 m/s, has a velocity at time t seconds given approximately by

$$v(t) = (20 - 10t)\ \text{m/s}$$

Determine

(i) the number of metres the ball is above the ground after 3 seconds;
(ii) the number of metres that it has travelled in that time.

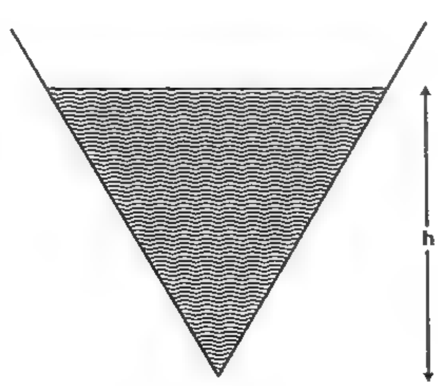
(Exercise 9.2.2.2, parts (iv) and (v), page 33, will give you a hint how to approach this problem if you are in difficulty.) ■

9.3.4 Some Other Applications

9.3.4 **Discussion**

So far in section 9.3 we have looked at the applications of definite integrals to volumes of solids of revolution, averages, velocities and distances. Two engineering applications are very briefly outlined below. If you are not familiar with them, pass them by; they are merely introduced for the benefit of those who understand their background.

Calculation of the total pressure on any surface immersed in water, such as a lock-gate or a dam, requires an integration, over the appropriate area, of the pressure at each point. The pressure in water increases linearly with depth. If a dam were triangular, this would lead to a definite integral of the form



$$\int_0^h (z \longmapsto z^2)$$

where z is the variable representing the depth below surface.

To determine the output of, say, a steam engine requires the evaluation of a definite integral of the type

$$\int_{v_1}^{v_2} (f : v \longmapsto p)$$

where v is the volume, p the pressure in the cylinder and f is a known function which maps v to p.

There are many more examples of the use of the definite integral in science and engineering, and in your further studies you will doubtless come across some of them. It is also interesting to note that a certain electrical process can be represented by a definite integral and that this fact can be used to advantage in the reverse sense: that is, we can design an electrical instrument which will evaluate definite integrals. For the moment we will use the word integrate to mean evaluate a definite integral. It will have a wider meaning in *Unit 13, Integration II*, but this one will suffice for the present. There is also a mechanical instrument, called a planimeter, with which we trace out the boundary of the required area and which then records this area. If you think about it, a rain-gauge is an integrator. It is even an integrator if you don't think about it*! It integrates the rate of rainfall over a specified period. And so also for that matter is a barometer — it adds up, or integrates, the weights of all the air molecules above it, the density (= mass/volume) at each level in the atmosphere being represented by quite a complicated function.

Definition 1

Finally, look up the definition of "integral" in a dictionary. It is one of the words which is given a special mathematical meaning and yet retains much of its everyday sense.

9.4 MORE REFINED APPROXIMATION METHODS FOR DEFINITE INTEGRALS

9.4

9.4.0 Introduction

9.4.0

Introduction

In this unit we have defined the definite integral of a function as the limit of a sequence, and we have shown that for certain cases (simple polynomial functions) we can evaluate this limit. In *Unit 13, Integration II*, we shall obtain a general result which will allow us to extend the set of functions for which this evaluation is practicable. But even then there will still remain many functions for which that general process is impracticable, for example,

$$\int_0^1 x \longmapsto \frac{1}{\sqrt{x^3+1}},$$

or for which it is unnecessarily complicated, for example,

$$\int_0^{0.5} x \longmapsto \frac{x^5}{\sqrt{1-x^2}}$$

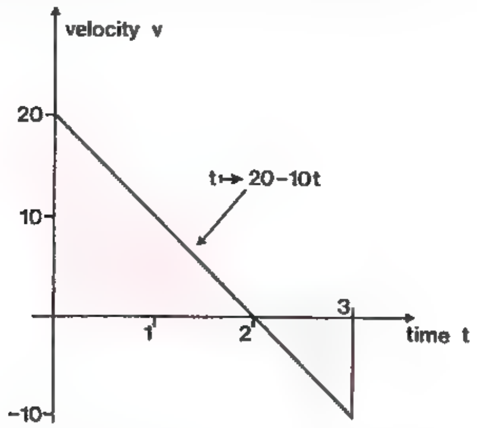
When we meet integrals such as these in practical work, we usually need a numerical answer correct to some given accuracy, so we can often usefully return to our original approximation processes (or variations of these). Thus we obtain an *estimate* of the limit to a given accuracy, rather than an exact formula for the limit.

* We add this comment to our colloquialism because our next unit is on logic!

Solution 9.3.3.1

Solution 9.3.3.1

Distances are represented by areas on the following diagram, as in the previous diagram for the simple example of the hiker.

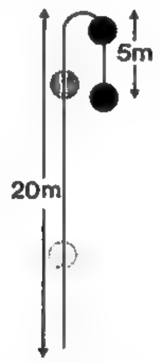


We could in this case use the formula for the area of a triangle, but we will apply integration techniques to illustrate the general method.

The graph crosses the t-axis at $t = 2$, indicating that after 2 seconds the velocity of the ball is zero. Subsequently the velocity is negative, i.e. the ball is returning to earth, so the ball will be at its highest point after 2 seconds. To calculate the distance travelled up to this time we evaluate

$$\int_0^2 t \longmapsto (20 - 10t) = 20$$

So the ball is 20 m above its starting point after 2 seconds.



In the next second it travels a distance given by

$$\int_2^3 t \longmapsto (20 - 10t) = -5$$

the negative sign indicating, as expected, that the ball is returning towards the ground during this period.

The answers are, therefore:

(i) $(20 - 5)\,\text{m} = 15\,\text{m}$,
(ii) $(20 + 5)\,\text{m} = 25\,\text{m}$.

The answer to (i), 15 m, is the result of the evaluation of

$$\int_0^3 (t \longmapsto (20 - 10t))$$

that is, the definite integral represents the distance from the starting point. Notice the more general point here that, if we form the definite integral over a domain for which the image changes sign, we need not split the domain up into sub-domains in which the numerical value is wholly positive and wholly negative. We need only do this in cases where the physical requirements of the problem demand it, such as distance travelled in part (ii) of the example above, or as in part (v) of Exercise 9.2.2.2. ■

9.4.1 The Trapezoidal Rule

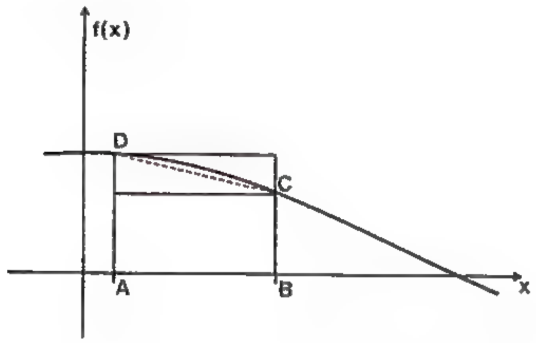
9.4.1

Main Text

We return now to a copy of the figure we used in the example at the beginning of section 9.1.2, when we were finding the area of part of a parabola. You will remember that we found a good estimate of the area by using

$$\tfrac{1}{2}(A_n + a_n)$$

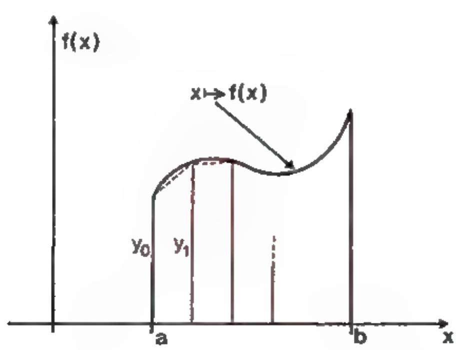
where $A_n =$ the sum of the n areas of the larger rectangles, and $a_n =$ the sum of the n areas of the smaller rectangles where $n =$ the number of intervals.



What does this mean geometrically? Half the sum of one larger rectangle and one smaller rectangle (an example is outlined in the figure) equals the area of the trapezium $ABCD$ (with upper boundary, CD, shown by the dashed line in the diagram). Thus the approximation we made earlier was equivalent to drawing the set of dashed lines (one for each interval) as the upper boundary to the area. That is, we approximate to the total area by the sum of the areas of the trapezia thus constructed.

We now turn to the case where $f(x)$ is non-negative in $[a, b]$.

Suppose we wish to find $\int_a^b f$ where the graph of f is given below:



We construct the set of dashed lines just as in the case of the parabola.

Suppose the ordinates of points on the graph at

$$a, a + h, a + 2h, \ldots, a + nh = b$$

are

$$y_0, y_1, y_2, \ldots, y_n \quad \text{respectively}$$

where

$$h = \frac{b - a}{n}$$

We have, from Exercise 9.1.0.1 on page 3:

(area of first trapezium on left) $= \frac{1}{2}(y_0 + y_1) \times h$

(area of second trapezium from left) $= \frac{1}{2}(y_1 + y_2) \times h$

and so on, until

(area of last trapezium) $= \frac{1}{2}(y_{n-1} + y_n) \times h$

Therefore, adding these equations together, we get:

the total area of all the trapezia

$$= \frac{h}{2}(y_0 + 2y_1 + 2y_2 + \cdots + 2y_{n-1} + y_n)$$

and therefore

$$\int_a^b f \simeq \frac{h}{2}(y_0 + 2y_1 + 2y_2 + \cdots + 2y_{n-1} + y_n)$$

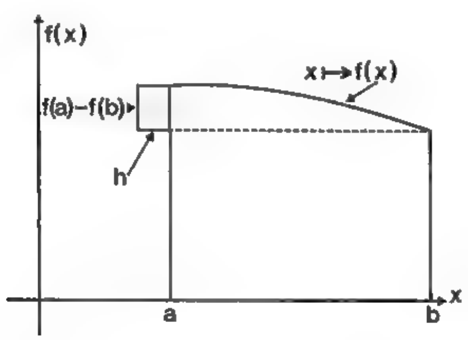
where $\simeq$ means "is approximately equal to".

This is known as the trapezoidal rule for evaluating definite integrals.* **Definition 1**

As indicated in *Unit 2, Errors and Accuracy*, we must now ask the question: "How approximate is 'approximate'?"

To answer this question we again consider the parabola. We use the fact that, for the rectangle approximation, provided the graph of the function slopes downwards over the whole interval or upwards over the whole interval, the difference, $A_n - a_n$, is the area of the largest rectangle. (See page 16.) In the case of the parabola:

$$A_n - a_n = h|f(b) - f(a)| \quad \text{where } h = \frac{b-a}{n}$$



The absolute error bound of the best estimate of the area, $\frac{1}{2}(A_n + a_n)$, is $\frac{1}{2}(A_n - a_n)$ which is

$$\frac{h}{2}|f(b) - f(a)| \quad \text{in the case of the parabola.}$$

The thing to notice about this formula is that, since $f(b)$ and $f(a)$ do not depend on the number of rectangles we choose to take in our approximation, the absolute error bound is proportional to h, the width of each sub-interval.

We return to the general case. The difference between this case and that of the parabola is that the graph of the function f no longer slopes downwards over the whole interval or upwards over the whole interval. (See the second figure on page 49.) However, we have only to split $[a, b]$ into smaller intervals in each of which the slope has the same sign throughout, and to apply the above argument to each of the smaller intervals. Since the "rectangle" method is equivalent to our trapezoidal

* See note on page 54.

rule, this means that the best we can say at this stage about the absolute error bound for the trapezoidal rule is that it too is proportional to h. In fact, it is usually more accurate than the rectangle method.

Example 1

Example 1

How many intervals would you need to take, *at most*, to evaluate

$$\int_0^1 x \longmapsto \frac{1}{x^3+1}$$

to an accuracy of two decimal places using the trapezoidal rule?

In practice, when evaluating $\frac{1}{x^3+1}$, we would round off the image values to some convenient number of decimal places, so that the error bound on each image will be ε, say. To how many decimal places must the images be calculated to ensure that the integral is accurate to two decimal places, and how many intervals will then be needed to ensure that the overall accuracy is to two decimal places? ■

Solution of Example 1

An absolute error bound using the trapezoidal rule is

$$\frac{h}{2}|f(1) - f(0)| = \frac{h}{2}|\tfrac{1}{2} - 1| = \frac{h}{4}$$

where

$$f: x \longmapsto \frac{1}{x^3+1} \qquad b = 1 \text{ and } a = 0$$

($f(x)$ decreases as x increases in [0, 1].)

For accuracy to 2 decimal places the absolute error bound must be less than or equal to $0.005 = 5 \times 10^{-3}$.

Therefore

$$\frac{h}{4} \leqslant 5 \times 10^{-3}$$

so

$$h \leqslant 2 \times 10^{-2}$$

and then, if the number of intervals is n, we have

$$n = \frac{b-a}{h} = \frac{1}{h} \geqslant 50$$

Therefore a minimum of 50 intervals is required to *guarantee* the required accuracy (with our present knowledge about the accuracy of the trapezoidal rule).

We have

$$\int_a^b f \simeq \frac{h}{2}\{y_0 + 2y_1 + \cdots + 2y_{n-1} + y_n\}$$

On the right-hand side there are $2n$ values of $f(x)$ inside the brackets. If each ordinate has an inherent error ε, the total error on the right-hand side will be (see *Unit 2, Errors and Accuracy*)

$$\frac{h}{2} \times 2n\varepsilon = nh\varepsilon = (b-a)\varepsilon$$

that is, a total error of ε in this particular case because $(b - a) = 1$ (note that this is independent of the number of intervals).

To be really sure that our *total* error from the use of the trapezoidal rule *and* the inexact data does not exceed 0.005, we can use 100 intervals (the absolute error introduced by the trapezoidal rule from this is then $\leqslant 0.0025$) together with data accurate to 3 decimal places (error from this is then $\leqslant 0.0005$) to give a total possible error of 0.0030; but in fact 50 intervals and data accurate to 3 decimal places would almost certainly suffice. ■

Exercise 1

Exercise 1
(5 minutes)

Repeat the last example with the definite integral

$$\int_0^2 x \longmapsto \exp(-x^2)$$

with a required accuracy of 3 decimal places. $[e^{-4} = 0.0183]$. ■

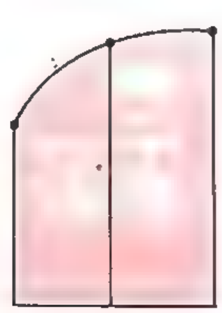
9.4.2 Simpson's Rule

9.4.2

Main Text

The trapezoidal rule gave an approximation to the area under a curve by using a set of straight line boundaries. One obvious way to improve the quality of the approximation is to take some account of the curvature of the boundary. In deriving the trapezoidal rule we took each pair of consecutive points $(a + mh, y_m)$ and $(a + (m+1)h, y_{m+1})$, $m = 0, \ldots, n-1$, found the straight line which passed through them (although we did not specify its equation since we already knew the area of a trapezium) and used this as the upper boundary of the area. We now introduce (no more than that) Simpson's rule for the area under a curve.

In *Unit 4*, *Finite Differences*, when we were studying polynomial interpolation (section 4.3.2), we found that we could often interpolate (i.e. find values between two tabular points) with increased accuracy if we increased the degree of the interpolating polynomial. We have an equivalent situation here. The linear interpolation rule uses the same degree of approximating polynomial, that is 1, as the trapezoidal rule (in which we replace the curved boundary by a set of line segments). The quadratic interpolation rule uses the same degree of approximating polynomial, that is 2, as Simpson's rule (in which we replace the curved boundary by parabolic segments). The basic element of area is now the one shown in the diagram covering two intervals with a parabolic upper bounding surface. Thus we now need to split the total interval into "2-interval" or "3-point" subsets.



What implication does this have on the number of sub-intervals we use? The answer to this is that the number of sub-intervals we use must now be even.

The bare outline of the basic steps in the argument are continued in the following text. You may like to check some of the steps and fill in the detail if you have the time; otherwise, read this section through quite quickly.

Consider a "3-point" subset in which the 3 points to be fitted on the given curve have co-ordinates $(-h, y_0)$, $(0, y_1)$, (h, y_2). The approximating quadratic polynomial function has the form

$$f: x \longmapsto a_2x^2 + a_1x + a_0 \qquad (x \in [-h, h])$$

It can be shown:

(i) by evaluating $\int_{-h}^{h} f$, that the area beneath the graph of f is

$$\frac{2a_2h^3}{3} + 2a_0h$$

(ii) by solving three simultaneous equations (or by using Lagrange's interpolation formula: see *Unit 4, Finite Differences*) that

$$a_0 = y_1$$

$$a_1 = \frac{y_2 - y_0}{2h}$$

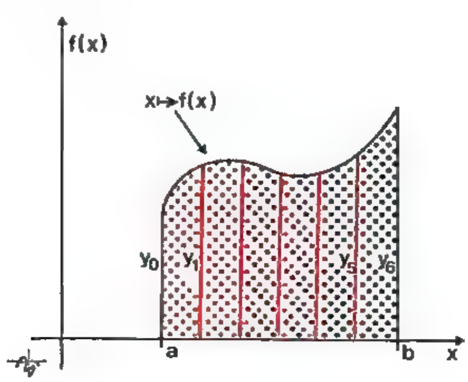
$$a_2 = \frac{y_0 - 2y_1 + y_2}{2h^2}$$

(iii) by substituting the appropriate results in (ii) into the result in (i) that the area beneath the graph of f is

$$\frac{h}{3}\{y_0 + 4y_1 + y_2\}$$

This result gives the approximation to the area beneath the original given curve in any "3-point" interval in which the ordinates are y_0, y_1 and y_2. Thus in the next "3-point" interval with ordinates y_2, y_3, y_4 (the interval in the figure shown dotted in black and red), the area approximation is

$$\frac{h}{3}\{y_2 + 4y_3 + y_4\}$$



Continuing in this way, we obtain the approximation (in the case of the function illustrated) to the total area as

$$\frac{h}{3}\{y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + 4y_5 + y_6\}$$

(continued on page 54)

Solution 9.4.1.1

Solution 9.4.1.1

Your answer should contain the following:

e^{-x^2} decreases as x increases in $[0, 2]$

$$n = \frac{2}{h} \geqslant \frac{1 - e^{-4}}{5 \times 10^{-4}} \simeq 2000 \text{ intervals}$$

If the error in each ordinate is ε, then the error in the result is 2ε.

So we could use 2500 intervals (producing a possible error of 4×10^{-4}) with data accurate to 4 decimal places (producing a possible error of 1×10^{-4}). ■

(*continued from page 53*)

In fact, the general formula with n intervals (n even) is

$$\int_a^b f \simeq \frac{h}{3}(y_0 + 4y_1 + 2y_2 + \cdots + 4y_{n-1} + y_n)$$

This formula is known as Simpson's rule.

Definition 1

It is generally true that you can obtain a better approximation (with the same number of intervals) using Simpson's rule than with the trapezoidal rule, but again we need calculus to demonstrate this more explicitly. You will probably think that if we replace the boundary in each interval by fitting a cubic to four points we will get a better approximation than with the quadratic fitted to three points and using Simpson's rule. This is not true, (although if we replace the boundary in each interval by fitting a quartic to five points, the accuracy does improve again). The trouble is that the complexity of the expression increases and one is in the usual "swings and roundabouts" situation. Only in cases when we get rather simpler formulation, such as when we use Gauss's eleven point formula in which a tenth degree curve is used to approximate to the boundary in each interval, does the corresponding formula get a name, and even then it tends not to be used because the modern computer can perform the simpler repetitive numerical procedure of the trapezoidal and Simpson's rule to the accuracy desired, by using many sub-intervals.

Note

We have obtained Simpson's rule and the trapezoidal rule for $\int_a^b f$ by considering the area beneath the graph of f between a and b. For simplicity, we have considered the special case in which $f(x) > 0$ for all x in $[a, b]$. In fact, these rules apply when $f(x)$ is not always positive in $[a, b]$. This can be seen by again considering areas (taking care of the signs) and slightly modifying our derivations.

9.5 CONCLUSION

9.5 Conclusion

In this text we have used the historical link between area and integration in order to develop the definition of a definite integral as the limit of a particular sequence of sums. We have seen some of the wider applications of the definite integral (to volumes of revolution and averages, for example) and have found specific formulas for the definite integrals of some simple polynomial functions. If the function to be integrated is not a simple polynomial function, the definite integral (if it exists) can still be evaluated by

either

(i) using the more sophisticated specific formulas which will be derived in *Unit 13, Integration II*, if these are appropriate

or

(ii) finding an approximate value by the powerful tools introduced in the last section of this text.

Acknowledgements

Grateful acknowledgement is made to the Cement and Concrete Association for permission to use the drawing of the Dollan Baths which appears on page 13, and also to the Mansell Collection for the portrait of Archimedes which appears on page 2.

Unit No.		Title of Text
1		Functions
2		Errors and Accuracy
3		Operations and Morphisms
4		Finite Differences
5	NO TEXT	
6		Inequalities
7		Sequences and Limits I
8		Computing I
9		Integration I
10	NO TEXT	
11		Logic I — Boolean Algebra
12		Differentiation I
13		Integration II
14		Sequences and Limits II
15		Differentiation II
16		Probability and Statistics I
17		Logic II — Proof
18		Probability and Statistics II
19		Relations
20		Computing II
21		Probability and Statistics III
22		Linear Algebra I
23		Linear Algebra II
24		Differential Equations I
25	NO TEXT	
26		Linear Algebra III
27		Complex Numbers I
28		Linear Algebra IV
29		Complex Numbers II
30		Groups I
31		Differential Equations II
32	NO TEXT	
33		Groups II
34		Number Systems
35		Topology
36		Mathematical Structures